

Student Growth in Elementary Mathematics:

A Cross Level Investigation

by

Seong Hee Kim

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved July 2012 by the
Graduate Supervisory Committee:

Finbarr Sloane, Chair
James Middleton
Alfinio Flores

ARIZONA STATE UNIVERSITY

August 2012

ABSTRACT

The primary purpose of this study is to examine the effect of knowledge for teaching mathematics and teaching practice on student mathematics achievement growth. Thirty two teachers and 299 fourth grade students in three elementary schools from one school district in urban area participated in the study. Most of them are Hispanic in origin and about forty percent is English Language Learners (ELLs). The two level Hierarchical Linear Model (HLM) was used to investigate repeated measures of teaching practice measured by Classroom Assessment Scoring System (CLASS) instrument. Also, linear regression and a multiple regression to examine the relationship between teacher knowledge measured by Learning for Mathematics Teaching (LMT) and Developing Mathematical Ideas (DMI) items and teaching practice were employed. In addition, a three level HLM was employed to analyze repeated measures of student mathematics achievement measured by Arizona Assessment Consortium (AzAC) instruments. Results showed that overall teaching practice did not change weekly although teachers' emotional support for their students improved by week. Furthermore, a statistically significant relationship between teacher knowledge and teaching practice was not found. In terms of student learning, ELLs have significantly lower initial status in mathematics achievement than non-ELLs, as were growth rates for these two groups. Lastly, teaching practice significantly predicted students' monthly mathematics achievement growth but teacher knowledge did not. The findings suggest that school systems and education policy makers need to provide teachers with the chance to reflect on their teaching and

change it within themselves in order to better support student mathematics learning.

DEDICATION

To my father.

ACKNOWLEDGMENTS

First, I would like to acknowledge my committee. I especially appreciate Dr. Sloane, my advisor, for his encouragement and support. He gave me a chance to work in this project for four years and to develop the main idea of my dissertation. Also, he gave me an opportunity to study methodology from many different perspectives. Whenever I was stuck, he helped me to better understand the problem and provided clues so I could solve it. From the beginning of my journey to this moment in graduate school, he guided me to keep going; sometimes fast and sometimes slow but never to stop. Dr. Sloane gave me intellectual and emotional support throughout this journey. One unforgettable issue happened to me, which is that he found me when I was lost on the highway on the way home from his house on Christmas day 4 years ago. I will never forget that Christmas! He was the first visitor, when I was in the hospital to give a birth of my daughter. He was also in the hospital with me and my husband, when my daughter needed emergency surgery. I would like to say “THANK YOU, Dr. Sloane!” When he reads this he may say that “you always say thank you,” but it has been sincere every time. I also appreciate Dr. Middleton. He was willing to discuss this work whenever I struggled. His support guided me to deeply think about the topic of my dissertation. I also appreciate Dr. Flores. His comments made me develop this study more thoroughly.

I would like to thank all the teachers and students who participated in this project. While I was working with them, I have learned about the system of the elementary school in United States. The teachers and students did not complain

regarding my poor English. Rather, they welcomed me to participate in their mathematics classes. Without their help, this work could not have been done.

I have wonderful colleagues. Thank you HyunJeong for taking care of me. She is another elder sister of mine in Arizona. Thank you Jennifer. She helped me to clean and organize the data and understand it more clearly. Thank you Brandon. During graduate school, he read my papers and gave valuable comments and English corrections. Thank you Meg. The conversation with her helped me to think through the findings of this study in plain English not just statistical numbers.

I appreciate my family's support for me. I have always felt sorry for them not being able to participate in any of family event. I would like to give special thanks to my elder sister, Seong Suk. She took care of me when I gave a birth and she worried more about my finishing the dissertation than I did. Her first visit to the US was more about staying in the apartment with baby but she never complained. I appreciate my mother and my younger sister. They have encouraged and supported me to keep working on my study. Last, I thank my husband, Kiwan and my daughter, Kaylee. Kiwan, he is my closest peer for my study. Sharing ideas with him and his comments helped me develop my ideas clearly. Since he was taking care of Kaylee while I was working on my dissertation, I now finish my PhD degree. Kaylee! Now, she is only four months old. As a first time mom and because of writing this dissertation, I was struggling but she has grown up well. "Thank you, Kaylee for your patience. Thank you, Kaylee for giving me such a beautiful experience as your mom. I LOVE YOU!"

TABLE OF CONTENTS

	Page
LIST OF TABLES.....	x
LIST OF FIGURES	xii
CHAPTER	
1 INTRODUCTION.....	1
Purpose of the Study	4
Significance of the Study.....	5
Dissertation Overview	6
2 LITERATURE REVIEW	7
Teacher Knowledge.....	8
Teacher Knowledge and Student Achievement.....	11
Teaching Practice.....	13
Teaching Practice and Student Achievement	15
Teacher Knowledge and Teaching Practice.....	17
Teacher Knowledge, Teaching Practice, and Student Achievement	20
Issues in the Effect of Teachers on Student Achievement Literature	23
Summary	26
3 METHODS.....	28
Sample	28
Data Collection	31

CHAPTER	Page
Student Data.....	31
Student Mathematics Achievement	31
Student English Proficiency Level.....	32
Teacher Data	32
Knowledge for Teaching Mathematics	32
Multiple Choice Items	33
Open Ended Items	34
Teaching Practice.....	35
Classroom Information	37
Procedure	38
Student Data.....	38
Student Mathematics Achievement	38
Student English Proficiency Level.....	39
Teacher Data	39
Knowledge for Teaching Mathematics	39
Teaching Practice.....	39
Classroom Information	40
Data Analysis	40
Description of Variables.....	40
Student Level Variables	40
Teacher Level Variables.....	43
Statistical Procedure	46

CHAPTER	Page
Summary	53
4 RESULTS	54
Examining Student Growth Graphically	54
Classroom by Classroom Analysis	57
Student Mathematics Achievement Growth	62
Change of Teaching Practice	68
Relationship between Knowledge for Teaching Mathematics and Teaching Practice	74
Effect of Teacher Knowledge and Teaching Practice on Student Mathematics Achievement Growth	76
Summary	81
5 DISCUSSION AND CONCLUSTION	84
Discussion	85
Student Mathematics Achievement Growth	85
Change of Teaching Practice	86
Relationship between Knowledge for Teaching Mathematics and Teaching Practice	88
Effect of Teacher Knowledge and Teaching Practice on Student Mathematics Achievement Growth	91
Policy Implications	95
Study Limitations	97
Recommendations for Future Study	99

CHAPTER	Page
Conclusion	103
REFERENCES	104
APPENDIX	
A UIVERSITY HUMAN SUBJECTS INSTITUTIONAL REVIEW	
BOARD	111
B SAMPLE DAP TEST ITEMS	113

LIST OF TABLES

Table	Page
1. Description of Demographic of Participant Fourth Grade Students ..	29
2. Grade of Participant Teachers.....	30
3. Ten Dimensions of CLASS measures	36
4. Means and Standard Deviations for MATHEMATICS	42
5. Description of CLASS scores Used for Research Questions	44
6. Means and Standard Deviations for the CLASS scores.....	45
7. Means and Standard Deviations for the TKNOWLEDGE.....	46
8. Fixed Effect at Three-Level Analysis of Mathematics Score on DAP Test with Unconditional Model	64
9. Random Effect at Three-Level Analysis of Mathematics Score on DAP Test with Unconditional Model	65
10. Variance Within and Between Classrooms at Three-Level Analysis of Mathematics Score on DAP Test with Unconditional Model..	66
11. Variance-Covariance Components and Correlation among the Level- 2 and Level-3 Random Effects at Three-Level Analysis of Mathematics Score on DAP Test with Unconditional Model..	67
12. Reliability of Level 1 and Level 2 Random Coefficient at Three-Level Analysis of Mathematics Score on DAP Test with Unconditional Model	67
13. Fixed Effect at Two-Level Analysis of PRACTICE with Unconditional Model.....	69

Table	Page
14. Random effect at Two-Level Analysis of PRACTICE with Unconditional Model.....	69
15. Reliability of Coefficient Estimate at Two-Level Analysis of PRACTICE with Unconditional Model.....	70
16. Fixed Effect Estimates at Two-Level Analysis of Teachers' overall CLASS and its Sub-domain Measures from Four Conditional Models	73
17. Random Effect Estimates at Two-Level Analysis of Teachers' overall CLASS and its Sub-domain Measures from Four Conditional Models	74
18. The Bivariate and Partial Correlations of the CLASS Sub-Domains and Teacher Knowledge Measure.....	76
19. Fixed Effect of Student ELL Status, Teacher Knowledge for Mathematics Teaching, and CLASS Scores on Student Mathematics Learning.....	79
20. Variance Decomposition from a Three-Level Analysis of the Effects of Student ELL Status, Teacher Knowledge for Teaching, and CLASS Measure on Student Learning	81

LIST OF FIGURES

Figure		Page
1.	Video Teachers' Rnking on MKT and Overall Lesson Scores	18
2.	Example of an LMT Item (KSC).....	34
3.	Student Mathematics Learning for All Participants.....	56
4.	Student Mathematics Learning for Each Classroom.....	58

Chapter 1

INTRODUCTION

Improving students' mathematics learning is the most basic goal of mathematics education. To understand and improve student learning, we study and develop educational interventions that can involve, but are not limited to, teacher professional development, the implementation of new curricula, the development of new assessments and their use in classrooms.

Among all those that influence student learning, teachers are considered the most important policy tractable factor (Wright, Horn, & Sanders, 1997). Although there are several such factors including family background (socioeconomic status, culture, or parents education level), and school factors (class size or school size), the teacher is considered the one factor that can be brought to bear on and effect on student learning (Wright et al., 1997). Put simply, teachers matter for student learning (Aaronson, Barrow, & Sander, 2007; Rockoff, 2004). To study teacher effects, researchers have examined certification, years of teaching experience, degree status, undergraduate or graduate courses taken, and major as teacher variables (Clotfelter, Ladd, & Vigdor, 2006; Goldenhaber, & Brewer, 2000; Monk, 1994). The results from studies employing these variables as predictors have been inconsistent. In addition, Hill, Rowan, and Ball (2005) concluded that this set of variables do not predict student' mathematics achievement. In light of the fact that teachers are seen as a critical policy lever, the question becomes what teacher variables can be shown to positively influence student mathematics achievement and growth.

A growing body of research has focused on teacher knowledge as the critical teacher factors needed to examine teaching and the variability of student achievement. Researchers found that teachers' knowledge significantly predicted student achievement (Carpenter, Fennema, Peterson, & Carey, 1988; Hill et al., 2005; Mullen, Murnane, & Willett, 1996). Further, teaching practice influenced student achievement (Cohen, & Hill, 2001; Fennema, et al., 1996; Hiebert, & Wearne, 1993; Silver, & Stein, 1996; Wenglinsky, 2002). A small body of research, however, has studied the effect of both teacher knowledge and teaching practice on student mathematics learning (Carpenter, Fennema, Peterson, Chiang, & Loef, 1989; Hill, Ball, Blunk, Goffney, & Rowan, 2007; Peterson, Carpenter, & Fennema, 1989; Rowan, Chiang, & Miller, 1997; Baumert et al., 2010).

Measurement becomes a fundamental issue in the study of the effect of teacher knowledge and teaching practice on student mathematics learning. Proxy variables (years of teaching, courses taken, etc.) for teacher knowledge do not well represent teachers' mathematics knowledge for teaching (Hill et al., 2005). Moreover, teachers' self-report from large-scale surveys of teaching practice can be far removed from what actually occurs in the classroom (Wenglinsky, 2002). Linking measures of teacher knowledge and teaching practice to student growth requires repeated measures of students over time (Sloane, 2003). Specifically, students should be measured on more than two occasions, and this has proven to be quite expensive. Finally, those wanting to analyze the effects of teachers must also consider the nested, or hierarchical, nature of the settings under study. Students in the same classroom are more likely to have similar learning

experiences when compared to those from other classrooms. This lack of statistical independence among nested subjects becomes critically important when we analyze data from real world settings including classrooms and schools.

In particular, measuring students' learning to evaluate change over time or to examine the quality of the educational system is difficult as learning is a process. In this study the rate of change in students' mathematics achievement is used as the proxy of learning. Sloane (2003) explained that "achievement is the sum-total of all the knowledge and skills possessed by an individual, and learning is the process by which knowledge and skills are acquired. Learning produces achievement, and as existing knowledge and skills are often frequently necessary to acquire more knowledge and skills, achievement affects learning" (p. 87).

Although learning and achievement are not exactly the same, achievement is the product of learning and measurable through testing. In this study, achievement represents learning at fixed points in time and is measured by students' scores from tests taken over time. Thus, the rate of change in achievement over time is then considered a measure of learning.

In summary then, policy researchers have long been interested in estimating teacher effects on student mathematics learning. In those studies, teacher quality was quantified as certification, major, degree, years of teaching experience, or number of courses taken. Although such characteristics are concrete, they do not explain teachers' mathematics classrooms where actual teaching and learning occurs. In this study, I examine the effects of teachers on student mathematics learning emphasizing the interactions between teacher

knowledge, as measured by Ball and Hill (2008) and Bell, Wilson, Higgins and McCoach (2010), and their teaching practices, as measured by Pianta, Laparo & Hambre (2008), in their mathematics classrooms. In doing so, I account for the nested structure of the data presented in real classrooms something rarely addressed by mathematics education scholars.

Purpose of the Study

Drawing on prior research, it seems reasonable to assume that teacher knowledge and teaching practice play an important role in student mathematics achievement. However, there is still no clear understanding of student growth in mathematics influenced by teachers. Thus, the primary purpose of this quantitative based correlational study is to estimate whether teachers' mathematics knowledge for teaching and their mathematics teaching practice predict fourth grade student mathematics growth (accounting for the nested structure of classroom learning environments). This study attends to the following research questions:

- (1) How does students' mathematics achievement change over the course of a school year?
- (2) Do teachers' instructional practices change over the course of a school year?
- (3) How well does mathematics knowledge for teaching predict teaching practice? And how well does teaching practice predict teacher's mathematics knowledge for teaching?

(4) Do teachers' mathematics knowledge for teaching and their teaching practice significantly predict students' mathematics achievement growth?

Significance of the Study

In this section I will discuss the significance of the study. First, I explain the significance of measurement issues; second, I consider the implications of the results to educational practice and policy decisions.

In terms of methodology, this study has two implications. First, concerning the measurement issues, teachers' mathematics knowledge for teaching is measured by a test. Teaching practice is measured through observations. Teacher knowledge and teaching practice are not collected from proxy variables or teachers' self-report. To measure these critical variables, instruments with validity and high reliability were adapted from other researchers' work. Further, student achievement is measured three times to understand and estimate their learning growth. Three times measures allow us to examine the linear growth. Second, Hierarchical Linear Model (HLM) is used to analyze the data set. Student achievement is repeated measures so each time measure is nested within an individual student. In addition, students are nested within classroom. Hence, the data were analyzed through a three level formulation. Further, teacher variables such as teacher knowledge and teaching practice are not separately analyzed. Rather, both teacher variables were analyzed in the same model simultaneously.

The results will help us to better understand change in teachers' practices over a school year, the relationship between teacher knowledge and teaching

practice, and the effect of teacher knowledge and teaching practice on the mathematics growth of their students. The findings can inform us in producing basic guidelines for the professional development in support of teacher learning. Further, knowledge of whether and how teachers influence student growth in mathematics can contribute to better educational policy decision making in support of teacher education and teacher assessment.

Dissertation Overview

This dissertation is organized as follows. In chapter one, I described the use of achievement in this study as representing learning. Further, I briefly outlined issues of research on teacher effects as they relate to student achievement. In chapter two, I focus on the review of the literature on teacher knowledge, teaching practice, and the effect of teacher knowledge and teaching practice on student mathematics achievement paying particular attention to the issues of measurement and analysis. In chapter three, I present a methodology of the current study including descriptions of statistical models. Along with a graphical depiction of classroom learning in mathematics, I present the quantitative findings in chapter four. Lastly, I provide a discussion drawing on the findings discussing the conclusions and their policy implications.

Chapter 2

LITERATURE REVIEW

Research demonstrates that teachers play the most critical role in student learning of mathematics. Consequently, teacher knowledge and teaching practice combine as the main factors of teacher quality known to enhance student mathematics achievement. The primary goal of this literature review is to examine the research literature in this domain of teacher knowledge and teaching practice, and critically evaluate them and their relationships to student mathematics achievement focusing on the measurement and analysis issues.

This literature review is divided into six sections. A brief description of each section follows. I then close the chapter with a short summary. The first section of this chapter offers an analysis of research on teacher knowledge focusing on the results of each study and the measures of teacher knowledge. In the second section, I provide the analysis of research on the relationship between teacher knowledge and students' mathematics achievement. The third section analyzes research on teaching practice, and the fourth section analyzes research on the relationship between teaching practice and students' mathematics achievement. The fifth section offers an analysis of research on the relationship between teacher knowledge and teaching practice, and the sixth section provides the analysis of research on teacher knowledge, teaching practice and student achievement, emphasizing the designs and analyses employed in these studies. In the final section, I consider the issues central to the study of the relationships

between and among teacher knowledge, teaching practice and student mathematics achievement.

Teacher Knowledge

Teacher knowledge plays an essential role in teaching and learning mathematics. A growing body of research has focused on teacher knowledge. Researchers have defined, categorized, and analyzed teacher knowledge. Shulman (1986) analyzed teacher content knowledge. In his results, he constructed three categories for such knowledge. These include subject matter knowledge, pedagogical content knowledge, and curricular knowledge (Shulman, 1986). Later, Ball and her colleagues reframed Shulman's definition of teacher knowledge to make it measurable (Ball, Thames, & Phelps, 2008; Hill, Schilling, & Ball, 2004; Hill, Ball, & Schilling, 2008). They argued for more specificity in the first two categories articulated by Shulman. They contend then that subject matter knowledge consists of common content knowledge, specialized content knowledge and horizon content knowledge. Further, they present pedagogical content knowledge as consisting of knowledge of content and students, knowledge of content and teaching, and knowledge of content and curriculum (Ball et al., 2008; Hill et al., 2008).

It is difficult to summarize simply the research on teacher knowledge because the pertinent researchers do not share a common focus. Consequently, their results can be seen as divergent. In this short summary I examine studies of professional development focusing on teacher knowledge. These research groups attended to different component of teacher knowledge to teach. Carpenter,

Fennema and Franke (1996) studied the impact of Cognitive Guided Instruction (CGI) on teacher knowledge of student strategies in whole number problem solving by contrasting the knowledge of CGI treatment teachers and their peers in the control settings. In addition, Hill and Ball (2004) focused on measuring teacher knowledge for teaching mathematics as the effect of Mathematics Professional Development Institutes (MPDIs) and showed teacher knowledge was improved by participating in MPDIs. Bell et al. (2010) compared teachers' specialized knowledge for teaching mathematics between Developing Mathematics Ideas (DMI) group and comparison group teachers and noted that DMI treatment teachers exceeded control group teachers' performance on the knowledge tests using multiple-choice items and open-ended items.

Other researchers again attend to different components of teacher knowledge to teach. Leinhardt and Smith (1985) studied differences in knowledge about fractions between novice and expert teachers. They found (1) differences in teachers' presentation (with considerable difference in the amount and level of conceptual information); (2) that expert teachers emphasized the role and value of algorithmic information differently (see p. 268 & 269). Ma (1999) compared knowledge of Chinese teachers and the U.S. teachers. She found that the Chinese teachers in the sample had a richer understanding of conceptual underpinnings of elementary number systems including whole numbers, the integers and the rational number. Finally, researchers affiliated with the Rational number project looked only at teacher content knowledge (Post, Harel, Behr, &

Lesh, 1988) and found teachers' narrow understanding of rational number concepts.

In summary, results from studies of teacher knowledge are both divergent and convergent. They are divergent in the sense that the researchers attended to different details. Some focused on mathematics content knowledge, others on knowledge to teach mathematics. Convergence occurs in that no matter the research focus, teachers with less experience (novice teachers) including preservice teachers look less prepared than expert teachers (Ball, 1990; Even, 1993; Graeber, Tirosh, & Glover, 1989; Tirosh, & Graeber, 1990; Simon, 1993). Moreover, U.S. teachers look less ready to teach mathematics than their China counterparts (Ma, 1999).

The tasks and methodologies associated with these studies also differ in their focus and attention to the details of teaching. Studies with a large number of teachers used a written test (Bell et al., 2010; Hill, & Ball, 2004; Hill et al., 2004). Ball (1990) and Ma (1999) interviewed participants. Most of the studies on preservice teachers (Even, 1993; Graeber, et al., 1989; Tirosh, & Graeber, 1990; Simon, 1993), and rational number project (Post et al., 1988) used written tests first for the all participants. They then selected some of the teachers and interviewed them. In particular, Leinhardt and Smith (1985) observed the classroom instruction of individual teachers and then interviewed them individually. The methods used to measure teacher knowledge interact with the results we observe. For example, multiple written tests allow us to compare teacher's knowledge in the group, interviews give us detailed information of what

teachers know about a concept, and observation of their instruction provides us insight into how teacher knowledge is used in classroom.

Teacher Knowledge and Student Achievement

An important consideration in improving student mathematics learning is to improve the quality of teacher knowledge of school mathematics (Ma, 1990). To advance this consideration, we must ask what the relationship between teacher knowledge and students' mathematics achievement is. Some researchers have focused on the relationship between teachers' mathematics knowledge and student mathematics achievement in an effort to answer the question. However, the results from those studies are not consistent with each other.

Some researchers found that teacher knowledge correlated with students' mathematics achievement. For example, Mullen et al. (1996) found that teacher mathematics knowledge was highly and positively related to student learning of advanced mathematics concepts. Hill et al. (2005) found that the only variable, in their study, to predict student mathematics gains over the 1st and 3rd grades was teachers' mathematics knowledge for teaching. Further, they found that courses taken, certification, years of teaching experience, and content knowledge for teaching-reading did not significantly predict students' mathematics learning. On the other hand, Carpenter and his colleagues (Carpenter et al., 1988) found that some of teacher knowledge is related with student achievement but some is not. They noted that teacher knowledge about problem type, problem difficulties, and the strategies that students used were not significantly correlated with student mathematical success. In contrast, teachers' prediction of student success in

problem solving was significantly correlated with student achievement.

Eisenberg (1977), however, noted that teachers' subject matter knowledge had little effect on student performance.

Many studies used proxy variables for teacher knowledge such as the number of mathematics/ mathematics education courses taken, mathematics major, years of teaching experience, mathematics master's degree, and certification. The findings from studies using proxy variables are inconsistent. Different measures of teacher knowledge produce varying results in the research literature.

Goldenhaber and Brewer (2000) found that differences in teacher certification do not predict student achievement. Yet, Clotfelter et al. (2006) found that variability in teachers' licensure test score significantly predicted their students' mathematics performance. Monk (1994) concluded that the number of mathematics/mathematics education courses taken by teachers in undergraduate, degree and major influenced student achievement differently. Moreover, he noted that their relationship was stronger for the advanced classes and less so for the remedial classes.

The results presented above are quite varied and inconsistent. This may be due to the fact that these researchers conceived their research questions, measures, and locations for study differently. Some researchers generated their own measures (Hill et al., 2005; Carpenter et al., 1988) and others used proxies for mathematics teachers' knowledge (Clotfelter et al., 2006; Goldenhaber, & Brewer, 2000; Monk, 1994). As consequence, those researchers do not define and

operationalize the act of teaching and teachers' knowledge for teaching in the same way.

Teaching Practice

Along with teacher knowledge, teaching practice is another important factor for teaching and learning mathematics. The teaching practice tells us what actually happens in mathematics classrooms where teaching and learning mathematics take place. The topics researchers examine when they consider teaching practice are varied. Through the analysis of teaching practice, researchers found different meaningful information. Cobb, Wood, Yackel, and McNeal (1992) compared and contrasted social interaction between traditional mathematics classroom and inquiry mathematics classroom. The meaning of success differs in the classroom that focuses on procedural knowledge and in the classroom that examine mathematics through inquiry. When students can follow procedural instructions successfully, it was meaningful in traditional mathematics classroom but when students can create and manipulate mathematical objects with their explanation and justification, it was meaningful in the inquiry mathematics.

Stigler and Hiebert (1999) compared teaching practice internationally. They analyzed the typical 8th grade mathematics lessons from Germany, Japan and the United States and found general patterns of teaching in each country that varied across countries. Stein, Smith, Henningsen, and Silver (2000) analyzed how the cognitive demand of mathematics tasks changed during instruction. When teachers set up tasks with high cognitive demand, several factors lowered it. These factors included becoming routinized task, teacher's emphasizing

correctness of the answer, not enough time, classroom management, inappropriateness of task, and teachers' no more accounting for high-level process. However, several other factors maintained high cognitive demand of the task including scaffolding, students' monitoring their own progress, model of high-level performance, press for high level process, tasks based on students' prior knowledge, teachers' conceptual connections, and sufficient time.

Other researchers were interested in mathematical discourse during instruction (Lampert, 1990, 2001; Williams, & Baxter, 1996). Lampert (2001) described and analyzed a variety of teaching practice such as establishing a classroom culture, preparing a class, and teaching while individual work and whole class from the beginning of school year to the end of school year. William and Baxter (1996) analyzed the patterns of discourses in mathematics classroom where the teacher supported producing mathematical knowledge through discourse among students. In the classroom students participated in the scaffolding of mathematical ideas and that of norms for social behavior and expectations as well. Cobb and his colleagues (McClain, & Cobb, 2001; Yackel, & Cobb, 1996) analyzed sociomathematical norms in the class. They found that teachers play a critical role in establishing norms for mathematical aspects of classroom activity and the mathematical quality in classroom. Raymond (1997) studied teaching practice to figure out its relationship with teacher belief. From six teachers, he noted that teachers' belief and their practice was inconsistent. The inconsistency came mainly from social teaching norms (school philosophy, administrators, standardized tests, curriculum, textbook, other teachers and

resources) and immediate classroom situation (students— ability, attitudes and behavior—, time, constraints, and the mathematics topic at hand).

It is interesting to note that the results of these studies varied while the methods engaged by these researchers appear consistent. The general method was that of observation. However, the duration of these data collection varied considerably with Lampert exploring her own practice over the course of a school year.

Teaching Practice and Student Achievement

This question, how does teaching influence student achievement, is essential to ask. Researchers have studied the relationship between teaching and student achievement. Hiebert and Wearne (1993) note that student achievement is generally higher in classrooms where, teachers focus on the following: (1) A small number of focused problems, in contrast to a lot of problems; (2) the problems are presented in story (or word) format; (3) students are allowed adequate time to engage the problems carefully; (4) where classroom and group discussion is valued; (5) where students are allowed and encouraged to build physical models; (6) where students orally defend their choices when solving the problem; and (7) where students produce more than a single solution strategy. Further, Stein and Silver (1996), and Wenglinksy (2002) note that the classroom where students have higher achievement also display the following characteristics: (1) teachers emphasize higher level thinking and use reasoning tasks; (2) teachers encourage students to come up with multiple solution strategies and representations; and (3) teachers ask students to explain their solution. In

contrast, students who have lower achievement work in classrooms where teachers emphasize procedural tasks, single solutions, single representations, and little mathematical communication between teacher and students, (Silver, & Stein, 1996).

Since teaching plays a critical role in student achievement, researchers have studied the relationship between teaching practice and student achievement to verify the effect of other educational product (e.g. professional development, and curriculum). For example, to estimate the effect of professional development, researchers have found that teachers' participating in professional development improved their instruction and the changed instruction highly influenced their students' mathematics performance (Cohen, & Hill, 2001; Fennema, et al., 1996). McCaffrey et al. (2001) studied teaching practice and student achievement to understand the effect of curriculum. They showed that students whose teachers engaged reform teaching practices in an integrated mathematics course had higher scores, in comparison to students whose teachers delivered traditional curriculum using reform strategies.

The results strongly indicate that teaching practice affects students' achievement in profound ways. In terms of data collection, these researchers used diverse methods. Some of them measured teaching practice by teachers' self-report (Cohen, & Hill, 2001; McCaffrey et al., 2001; Wenglinsky, 2002) and the other observed mathematics classrooms (Fennema, et al., 1996; Hiebert & Wearne, 1993; Silver, & Stein, 1996). To measure students' achievement, some developed their own test items (Fennema et al., 1996; Hiebert, & Wearne, 1993)

and others used district or NAEP assessments. While they measured students' achievement in various ways, they analyzed the achievement data in the similar way. Most studies used students' achievement as a single time-point dependent variable; one study used the previous year score as covariate (McCaffrey et al., 2001) and Fennema et al. (1996) analyzed the changes in class means. Moreover, the way in which researchers analyzed teaching practice and student achievement was very similar. Generally they analyzed teaching practice and student achievement separately and made the conclusions from the two findings. This was not the case of Fennema et al. (1996) who related instructional change to student achievement in the analysis.

Teacher Knowledge and Teaching Practice

Both teacher knowledge and teaching practices combine in the act of instruction in mathematics classrooms. Examining instruction raises pertinent questions about the combination of these factors, how those two are related with each other. Teaching is highly dependent on teachers' knowledge (Ball, 1991; Putnam, Heaton, Prawat, & Remillard, 1992). Hence, teachers' lack of knowledge influences their teaching. For example, Borko, Eisenhart, Brown, and Underhill (1992) found that a preservice teacher's lack of mathematical knowledge was reflected in her unsuccessful teaching of the concept of fraction division algorithm. In addition, others have seen that a teacher with limited subject matter knowledge led to the narrowing of instruction in three ways: "(a) the lack of provision of ground work for future learning in this area, (b) overemphasis of a limited truth, and (c) missed opportunities for fostering

meaningful connections between key concepts and representations” (Stein, Baxter, & Leinhart, 1990, p. 659).

More recently, researchers studied different levels of teacher knowledge and teaching practice. Hill and her colleagues (Hill et al., 2008) examined the link between teachers’ mathematical knowledge for teaching and the quality of their teaching practice with various levels of teaching practice and knowledge. Ten teachers in this study were ranked on teacher knowledge and teaching practice as can be seen in the figure below.

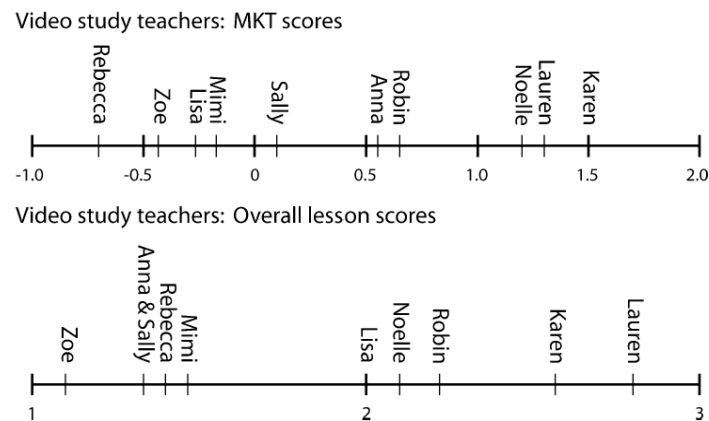


Figure 1. Video teachers' ranking on MKT and overall lesson scores. Adapted from “Mathematical Knowledge for Teaching and the Mathematical Quality of Instruction: An Exploratory Study,” by H. C. Hill, M. L. Blunk, C. Y. Charalambous, J. M. Lewis, G. C. Phelps, L. Sleep, and D. L. Ball, 2008, *Cognitive and Instruction*, 26, p. 444. Copyright 2008 by Taylor & Francis Group.

Hill et al. (2008) analyzed five teachers’ instruction to examine the relationship between mathematics knowledge for teaching and instruction. The five cases are Lauren (high knowledge /high instruction), Zoe (low knowledge/

low instruction), Anna (worse practice than knowledge), Rebecca (better instruction than knowledge), and Noelle (fifth high-rank case). From those cases, they found that what teachers know and how they teach is powerfully related. They argued that the variation in the quality of observed instruction was due to the differences in knowledge for teaching mathematics in their sample cases.

In particular, Kahan, Cooper and Betha (2003) compared two teachers' instructional action: one with high and the other with intermediate mathematical content knowledge. From the analysis of the two cases, they found that the scope of what was possible in mathematics class was constrained by teachers' limited mathematical content knowledge because "all these aspects of MCK (*mathematics content knowledge*) play a role in perceiving and seizing the teachable moment" (Kahan et al., 2003, p. 246). However, they also found that mathematical content knowledge alone did not guarantee exemplary teaching. In essence, mathematical content knowledge can be seen as a necessary condition for good instruction but it is unlikely to be a sufficient condition.

To measure teacher knowledge, researchers interviewed teachers or gave them a written test. Most studies but Hill et al. (2008) measured teachers' mathematics subject matter knowledge rather than knowledge for teaching mathematics. To measure teaching practice, all the studies observed mathematics classes. In analyzing teaching practice, researchers conducted case studies focusing on several selected teachers. Consequently, these researchers described the relationship between teachers' knowledge and teaching practice with rich

detail. In these studies the analytic lens used was that of the “case” study. It could be argued that we need studies of larger samples.

Teacher Knowledge, Teaching Practice, and Student Achievement

How do both teacher knowledge and teaching practice influence student mathematics achievement? A growing body of research has tried to examine this question. Researchers have collected data of teacher knowledge, teaching practice and student achievement. However, the way in which they examine the relationship between teacher data and student data varies. I found four different types of analysis from the literature.

The first type of analysis was to compare each of the three data set separately. Carpenter et al. (1989) compared teachers’ knowledge, instruction, and student mathematics achievement separately when comparing CGI classrooms and non-CGI classrooms. In terms of teacher knowledge, CGI teachers were significantly better in predicting student strategies used. For instruction, CGI teachers spent significantly more time on posing problems, solving word problems and listening to students’ explanations about the process. Further, they expected their students to use multiple strategies. On the other hand, control class teachers spent significantly more time on number fact problems and providing feedback to students’ answers, and they expected students to use counting strategies. In terms of students’ achievement, CGI students displayed no difference on computation test. However, they outperformed control group students in complex addition and subtraction problems.

In the second type of analysis, researchers examined how teacher knowledge and teaching practice separately influence student achievement. Rowan, Chiang, and Miller (1997) analyzed how teacher knowledge and teaching practice predict student achievement. For teachers' mathematics content knowledge, they collected a mathematics quiz score from NELS:88. They concluded that teacher mathematics content knowledge was highly related to student mathematics achievement. For teaching practice, they collected instructional strategies from five survey items and found that teaching practice did not predict student mathematics achievement. Their results may be highly dependent on the quality and meaning of their measures.

The third type of analysis was to focus on teacher knowledge and find the effect of teacher knowledge on teaching practice and student achievement. Hill and her colleagues (Hill et al., 2008) studied how teacher knowledge predicts teaching practice and student achievement using two different data sets. From the data of MPDIs, they found that teachers' high score on mathematics content knowledge was highly correlated with high scores on their classroom instruction. In addition, from the data in their previous study (Hill et al., 2005); they found that teachers' knowledge for teaching mathematics accounted for significant variability in student mathematics achievement. From the two studies, they concluded that teachers' content specific knowledge predicts their classroom instruction and students' achievement. The analytic style of Peterson and her colleagues' work (1989) is similar to that of Hill et al. (2008). Peterson et al. found that teachers' knowledge of students' problem-solving was positively

related to teachers' questions about the problem solving process and teachers listening to students' solutions during instruction. Moreover, teachers' knowledge of students' problem-solving was also positively related with students' problem solving achievement.

The fourth, and final, analytic frame identifies the effect of specific types of teacher knowledge on teaching practice and student achievement. Baumert et al. (2010) distinguished content knowledge and pedagogical content knowledge as separate component of teacher knowledge. They studied whether teachers' content knowledge or pedagogical content knowledge makes a unique contribution explaining the differences in the quality of instruction and student progress (Baumert, et al., 2010). They found that pedagogical content knowledge predicted teachers' quality of instruction and student progress powerfully. Content knowledge, however, is less predictive for student progress and has no direct effects on instructional quality, although content knowledge was highly correlated with pedagogical content knowledge. The structure of this study is similar to that of Peterson et al. (1989), which predicted the effect of teacher knowledge on teaching practice and student achievement. In contrast, Baumert and his colleagues studied both content knowledge and pedagogical content knowledge, whereas Peterson et al. studied only pedagogical content knowledge.

Each of the five studies measured teacher knowledge with different method and collected different type of teacher knowledge. Carpenter et al., (1989) and Peterson et al. (1989) interviewed teachers, and others used a paper and pencil test. Some measured pedagogical content knowledge (Carpenter et al,

1989; Peterson et al., 1989), another measured teachers' mathematics content knowledge (Rowan et al., 1997), and the others measured both pedagogical content knowledge and content knowledge (Hill et al., 2008; Baumert et al., 2010). We can see this as a growing sophistication over time. Earlier studies treated the problem with less complexity than the later studies have.

Many studies observed mathematics classes to measure teaching practice except Rowan et al. (2009) and Baumert et al. (2010). Rowan et al. (1997) used data from surveys which draw on teachers' self reports. Baumert et al. (2010) collected tests and examinations used in the class and analyzed the cognitive demand of the tasks for teaching practice. They also collected students' rating scales about teachers' support for their learning.

Studies collected different types of students' achievement data and analyzed them in various ways, and from various perspectives. Carpenter et al. (1989) and Peterson et al. (1989) used the Iowa test and also developed their own items. Rowan et al. (1997) used NELS data, and Baumert et al. (2010) used the federal test data. One used the student achievement score as a single time-point dependent variable (Baumert et al., 2010); while another used the previous year's score as covariate (Rowan et al., 1997). The others used pre-post test differences as the dependent variable (Carpenter, et al., 1989; Hill et al., 2008; Peterson et al., 1989).

Issues in the Effect of Teachers on Student Achievement Literature

In this section I will critique the literature on teacher effects on students' mathematics achievement based on the previous sections. First, I look at

measurement in the literature; second, I examine the unit of analysis; and third, I investigate the structure of the analysis.

To measure student achievement, many researchers developed their own test items. Some researchers used district test or state test; others used nationally representative assessment. These tools varied in quality (e.g., validity and reliability). Although they used different assessments, how they used the test results was similar. Most studies analyzed one-time measure as a dependent variable without following up the students' score. Generally these studies employed randomized trials or quasi-experiment. Occasionally, some researchers used the test score from the previous year as covariate (e.g., McCaffrey et al., 2001; Rowan et al., 1997) or pre-post difference (e.g., Carpenter et al., 1989; Hill et al., 2005; Mullens et al., 1996; Peterson et al., 1989) but those are very rare in the body of work.

For teacher knowledge, researchers measured teachers' pedagogical content knowledge or content knowledge. We know today that only subject matter knowledge is not sufficient for fully understanding teaching because the subject matter knowledge should be transformed into pedagogical content knowledge (Even, 1993; Shulman, 1986, 1987).

For the research on teacher knowledge itself, most studies measured teacher knowledge via written test or interview. For the research on the relationship between teacher knowledge and student achievement, however, many researchers employed proxy variables from national data sets or using scores from eighth grade mathematics tests or certification tests as their measure of teacher

content knowledge. The results from proxy variables for teacher knowledge were inconsistent across studies. This implies that the proxy variables substituted for teacher knowledge did not represent well teachers' mathematics knowledge for teaching (Hill et al., 2005). Besides, because teachers took the licensure tests before they became teachers, the scores could not reflect teachers' learning from their teaching experiences in their classrooms. Teachers' experiences from their own teaching need to be considered when measuring teacher knowledge because teachers learn about teaching mathematics while they are teaching (Franke, & Kazemi, 2001).

Measuring practice has similar issues to those seen with the measurement of teacher knowledge. When researchers examined teaching practice itself, most studies observed mathematics class, but when they examined the relationship between teaching practice and student achievement, many of them used teachers' self-report from surveys. This requires a rather large informational jump. Although data from large-scale surveys are easily collected, teachers can choose to report only idealized instruction, instead of their actual instruction so the results can be far removed from what actually occurs in their mathematics classroom (Wenglinsky, 2002). Therefore, observations are necessary to measure teaching practice.

Many studies analyzed student mathematics achievement scores either as aggregated across entire classes of individual students, which implies that classroom means were analyzed, or as disaggregated individual score ignoring the clustered characteristic of the classroom, which implies that individual scores

were analyzed. The achievement at classroom-levels did not necessarily reflect individual student achievement (Sloane, 2008a, 2008b). Although class means increased, many individual student achievements could have decreased or stayed at the same level if other students' achievements disproportionately increased. Moreover, students are nested in classrooms in a nonrandom fashion. Students in the same classroom have experienced the effects of working with a particular teacher over a prolonged period. Hence, students within a classroom are likely to be more similar than students from different classrooms. Therefore, careful data analysis must consider individual differences and classroom similarities simultaneously.

The majority of studies analyzed teacher data and student data separately and drew conclusions from the individual findings. Although the individual findings could be used to explain differences in knowledge, teaching practice and students' achievement between groups; sometimes as the result of an intervention, the researchers do not investigate the direct relations among them. No studies explicitly examined the linkages between teacher knowledge, the practice of teaching and student achievement. To study the effects of teachers on students' achievement, a different structure should be employed.

Summary

The studies of teacher effects on student mathematics achievement have measurement issues for teacher knowledge, teaching practice and student achievement, and analysis issues for the nested data set. By using repeated measures of student achievement, observing teaching practice and employing

multilevel analysis model, I addressed these issues in current study. In the next chapter, I propose a study that explores the correlation between teacher factors and student mathematics achievement growth.

Chapter 3

METHODS

This study examines the effects of teacher knowledge and teaching practice on student mathematics achievement growth. In this chapter I describe the overall methodology for the study including sample, data collection and procedure, variable description, and statistical procedure. Data are from teachers and students from three elementary schools. For the main research question, the effect of teacher knowledge and teaching practice on student mathematics achievement growth, a multilevel modeling with repeated measures was used. Level-1 is repeated measures of mathematics scores nested within the individual student, level-2 is student background information, and level-3 measures the impact of teacher knowledge and teaching practice. The purpose of this design is to consider the organizational setting where individual student scores are nested within each student and students are nested within teacher.

Sample

The sample in this study comprise all students and teachers at the fourth grade level in three elementary schools from one school district in urban area of a large city in the southwestern. Each school has a total of approximately 600 students and 30 teachers. Ninety-three percent of the students qualified for free lunch and the other seven percent qualified for reduced price lunch. Ninety-six percent of the children are Hispanic in origin, two percent of the children are African-American, and two percent of children are Caucasian. Almost half of the students in these schools are English Language Learners (ELLs). At these

schools, classes are organized by English proficiency level and teachers teach only in English. Because the three schools are all in the same district, students and school background are very similar. In particular, teachers in the three schools have received CGI professional development for five years.

For student data, 299 fourth grade students participated in the study for two years (127 from school year 2009 and 172 from school year 2010). Table 1 describes their demographic. Most of them are Hispanic in origin and about half is male and more than 40% is ELLs.

Table 1

Description of Demographic of Participant Fourth Grade Students

	School Year 2009 (<i>n</i> = 127)	School Year 2010 (<i>n</i> =172)	Total (<i>n</i> = 299)
Gender			
Male	70 (55%)	94 (55%)	164 (55%)
Female	57 (45%)	78 (45%)	135 (45%)
Ethnicity			
Hispanic	123 (97%)	172 (100%)	295 (99%)
Asian	1 (1%)	0 (0%)	1 (--)
Caucasian	2 (2%)	0 (0%)	2 (1%)
Indian	1 (1%)	0 (0%)	1 (--)
ELL status			
ELL	57 (45%)	68 (40%)	125 (42%)
Non-ELL	70 (55%)	104 (60%)	174 (58%)

Note. Percentages are in parentheses.

For teacher data, thirty two teachers participated in the study for two years (18 from school year 2009 and 14 from school year 2010). They taught second, third, fourth or fifth graders. Table 2 describes their grade level in detail.

Table 2

Grade of Participant Teachers

	School Year 2009 (n = 18)	School Year 2010 (n=14)	Total (n = 32)
Grade			
2 nd	4 (22%)	0 (0%)	4 (13%)
2 nd and 3 rd	0 (0%)	0 (0%)	0 (0%)
3 rd	4 (22%)	0 (0%)	4 (13%)
3 rd and 4 th	3 (17%)	1 (7%)	4 (13%)
4 th	4 (22%)	5 (36%)	9 (28%)
4 th and 5 th	0 (0%)	3 (21%)	3 (9%)
5 th	3 (17%)	3 (21%)	6 (19%)
5 th and 6 th	0 (0%)	2 (14%)	2 (6%)
Single-grade classroom	15 (83%)	8 (57%)	23 (72%)
Multi-grade classroom	3 (17%)	6 (43%)	9 (28%)

Note. Percentages are in parentheses.

Student data were analyzed to address research questions 1, student achievement growth, and question 4, the effect of teacher knowledge and teaching practice on student mathematics achievement growth. Teacher data were analyzed for research question 2, 3, and 4. For research question 2, which focuses on teaching practice, 14 teachers from school year 2010 participated because

teaching practice was repeatedly measured only in school year 2010. For research question 3, which attends to the relationship between knowledge and teaching practice, all 32 teachers participated. In addressing research question 4, sixteen fourth grade teachers participated.

Data Collection

Data for this study were collected from teachers and fourth grade students. For students, mathematics achievement scores and English proficiency level were collected. For teachers, their mathematics knowledge for teaching, and teaching practices were collected. Also, I collected whether he or she taught a multi-grade class or a single-grade class, and the English proficiency level of the class. I provide more information on the instruments used to collect teacher and student data below.

Student Data

Student mathematics achievement. Student mathematics achievement, measured in this study, focuses on the score a student acquired from a mathematics test. It was measured with the District Achievement Plan (DAP) using the Arizona Assessment Consortium (AzAC) instruments. Students take the DAP test four times in a school year with the same problem set but different order of the items. The test has about 70 items for each grade level. All students take the tests three times, and students who do not pass the standards take the test a fourth time. In this study I used data from the first, second and third testing only which all students have their score. In this school district, they have used the same DAP tests in each of three years. Thus the DAP test form from school year

2009 to 2011, employed the same items and this allowed us to combine data from the year 2009 and 2010. Consequently, I was able to generate a sample of sixteen classrooms for research question 1 and 4.

Student English proficiency level. Because almost half of the students are designated ELL, their ability to speak, comprehend, write and read in English is measured by their English proficiency level. Students' English proficiency level was measured with the Arizona English Language Learner Assessment (AZELLA). AZELLA, was developed by the Pearson Company and the Arizona Department of Education, is an augmented version of the Stanford English Language Proficiency (SELP) test. AZELLA includes forced-choice, semi-structured response, and unstructured response items. It is a criterion-referenced test. The results have oral, reading, writing and overall scores. Students take the AZELLA test at the end of school year and are categorized into ELL or Non-ELL categories. With the categorization comes classroom assignment. Students at the same English proficiency level are grouped in the same classroom.

Teacher Data

Knowledge for teaching mathematics. Teacher knowledge, as measured in this study, does not focus on general content knowledge but knowledge that teachers use in their instruction. Put differently, in this study, teachers' mathematics knowledge for teaching is not mathematics subject matter knowledge but mathematics pedagogical content knowledge. The test to measure teacher knowledge has 12 items adapted from Learning for Mathematics Teaching (LMT) (Hill et al., 2004), Developing Mathematical Ideas (DMI) (Bell et al.,

2010), and algebraic thinking from CGI (Carpenter, Fennema, Franke, Levi, & Empson, 1999). The test consisted of multiple choice-items and open ended-items.

Multiple choice items. There were seven multiple choice items. An example is shown in Figure 2. This example is from LMT. LMT sample items are online, http://sitemaker.umich.edu/lmt/files/LMT_sample_items.pdf. The multiple-choice items measure different dimensions of teacher mathematics knowledge for teaching. One item measured Knowledge of Content and Student (KCS), and the others measured Special Content Knowledge (SCK). KCS requires “knowledge of students and their ways of thinking about mathematics – typical errors, reasons for those errors, developmental sequences, and strategies for solving problems” (Hill et al., 2004, p. 17). Said differently, KCS is teachers’ knowledge of student mathematics understanding. SCK is “including building or examining alternative representation, providing explanation, and evaluating unconventional student methods” (Hill et al., 2004, p. 16). That is, SCK considers teachers’ ability to solve a problem and various ways of understanding the concept and the problem as well. The topics of multiple-choice items are about multidigit subtraction, place value, fraction, multiplying, and division in number and operations. Some items have sub-problems.

14. Mrs. Jackson is getting ready for the state assessment, and is planning mini-lessons for students around particular difficulties that they are having with subtracting from large whole numbers. To target her instruction more effectively, she wants to work with groups of students who are making the same kind of error, so she looks at a recent quiz to see what they tend to do. She sees the following three student mistakes:

I	II	III
$\begin{array}{r} 4 \ 12 \\ 502 \\ - 6 \\ \hline 406 \end{array}$	$\begin{array}{r} 4 \ 15 \\ 38008 \\ - 6 \\ \hline 34009 \end{array}$	$\begin{array}{r} 6 \ 9 \ 8 \ 15 \\ 7888 \\ - 7 \\ \hline 6988 \end{array}$

Which have the same kind of error? (Mark ONE answer.)

- a) I and II
- b) I and III
- c) II and III
- d) I, II, and III

Figure 2. Example of an LMT item (KSC). Adapted from Ball, D. L., & Hill, H.

(2008). *Mathematical knowledge for teaching (MKT) measures: Mathematics released items 2008*. Retrieved from

http://sitemaker.umich.edu/lmt/files/LMT_sample_items.pdf

Open ended items. There were five open-ended items. Those were drawn from DMI and CGI problems. DMI items measured SCK and KCS like the LMT items and also measured Knowledge of Content and Teaching (KCT).

KCT “combines knowing about teaching and knowing about mathematics” (Ball et al., 2008, p.401). Based on KCT, teachers generate content through a particular sequence and decide what mathematics material is needed to teach a concept (Bell et al., 2010). For example, teachers determine whether they give students a new problem or have more discussion on the previous problem. Further, teachers decide to use either fraction bars or pattern blocks to teach fractions based on their KCT. The DMI items had four stems that are subtraction,

multidigit multiplication, fair sharing, and fraction (Bell et al., 2010). The test included one item from each stem so four items are adapted from the DMI.

An algebraic thinking item from CGI work was then used to ask about strategies and answers students use to solve the relational thinking problem, $8 + 15 = \square + 16$. This item is thought to measure KCS.

Teaching practice. Teaching practice, as measured in the study, concentrated on the interaction between a teacher and students in mathematics class. Teaching practice is measured by Classroom Assessment Scoring System (CLASS) (LaParo, Pianta, & Stuhlman, 2004; Pianta, LaParo, & Hamre, 2008) rubric, a reliable measure for assessing classroom instruction (LoCasale-Crouch et al., 2007). The CLASS framework was developed to measure teaching quality in classroom contexts across grades and across content areas (Pianta, & Hamre, 2009). The instrument emphasizes what teachers and students are doing in the classroom rather than focuses on the materials they use in classroom or classroom environment (LaParo et al., 2004). The tool measures the interaction between teachers and students in three domains: emotional support, classroom organization, and instructional support. There are sub-dimensions for each domain. The following table (Table 3) shows the ten dimensions of the CLASS measures (Pianta et al., 2008).

Table 3

Ten Dimensions of CLASS measures

Domain	Classroom Dimensions
Emotional Support	(1) Emotional connection, respect, and enjoyment demonstrated between teachers and students, and among students; (2) Negativity such as anger, hostility, or aggression exhibited by teachers or students; (3) Teachers' awareness of and responsibility to student academic and emotional concerns; and (4) Classroom activities emphasizing on students' interests and motivations.
Classroom Organization	(1) The effectiveness of teacher monitoring, preventing, and redirecting students' behavior; (2) The productivity of using time to learn; and (3) Activities and materials to maximize students' opportunities to learn.
Instructional Support	(1) The use of instructional discussions and activities to promote students' higher-order thinking skills; (2) Teachers' extension of students' learning through teachers' responses to students' ideas, comments, and work; and (3) Teacher's encouragement of students' language.

The scores range from 1 (minimum score) to 7 (maximum score) for each dimension of classroom characteristics. The score 1 or 2 is low, moderate is from 3 to 5, and high is 6 and 7. Because there are 10 dimensions, the total score will range theoretically from 10 to 70 for an observation of a classroom using CLASS instrument.

Classroom information. Classroom information such as multi-grade classroom/single grade classroom and English proficiency level of the class was also collected. Students in a multi-grade classroom, for example 3rd grade and 4th grade combination classroom, would likely receive different curricular attention from their teacher when compared to students from a singly graded classroom because a teacher needs to cover grade level content for both grades at the same time. Arizona, like many other States, is undergoing severe fiscal problems. This has resulted in increased class sizes throughout the state and in the reduction of the workforce. Consequently, some of the classes in the study house students from one, two and three grade levels. From the classroom observations it was clear that teaching in multi-age, multi-grade classrooms is difficult for a number of the participating teachers. Consequently, I felt it important to control for this characteristic of the classroom learning environment and to estimate its impact.

All schools in the state of Arizona group their students by language proficiency. So this school district also grouped students by their English proficiency. If students are designated ELL, they are put in an English Language Development (ELD) classroom. Those students are committed to a 4-hour Language Arts block (reading, grammar, writing, and conversation). Students in

the proficient classes, however, are only committed to a 2-hour reading block which is mandated by the district, not the state. Although the school district requires a minimum one hour of mathematics a day, students from ELD classrooms and a proficient classroom are treated differently in the sense that they are not getting the same access to instruction. This policy is employed along with the within school tracking of students based of their ELL status. That is, classrooms are more homogeneous as students are assigned to classrooms based on the scores on the Arizona English Language Assessment (AZELLA). Thus, the ELL status of the classroom is an important one as students appear to be segregated within schools by first language competency. I use ELL_CLASS variable to help control for this unique difference between classrooms in the district under study.

Procedure

Data for this study were collected over two years. Student data and classroom information were collected from the school district. Teacher data were collected as the part of CGI professional development project. I provide more information on the procedures used to collect student and teacher data below.

Student Data

Student mathematics achievement. Students took DAP test in October, December, and March in school year 2009 and 2010. The raw score was collected to analyze. DAP scores were provided by the school district office.

Student English proficiency level. I collected if a student is ELL or non-ELL from the result of AZELLA test. It was collected from the school district office.

Teacher Data

Knowledge for teaching mathematics. All participant teachers took a test at the end of the school year (2010 and 2011). Because teachers learn how to teach mathematics, and how students think and solve mathematics problems while they teach mathematics in their class, they took the test at the end of the school year, after such learning. Teachers were encouraged to answer as much as they know for the open-ended questions. Enough time was allotted for teachers to complete most items. Two graduate students, including me, score teachers' tests for reliability between scores. Because LMT items are multiple choice items, the grade is correct or incorrect, and the score is the total number of correct ones. DMI items are graded based on "DMI assessment: Open-ended question scoring rubric". For a relational thinking item, the number of strategies teachers come up with of student strategies is their scores.

Teaching practice. To use the CLASS instrument, observers are required to complete the CLASS training to become certified as a reliable observer. This involves three days training from Teachstone (created by Bob Pianta and Bridget Hamre, who are CLASS authors). After having the training, the participant takes the CLASS reliability test to be certified to use CLASS measure to observe classroom interaction. To pass the test, the observer scores six classroom instruction videos which are about twenty minutes and the scores for each

dimension should be reliable. After certified as reliable, the data were collected. Six graduate students in school year 2009 and five graduate students in school year 2010, including me in both years, observed eighteen and fourteen classes more than four times for each class to prevent special case of instruction biasing teachers' scores over a school year. I collected the scores from ten dimensions in CLASS separately and overall scores from teachers. Every effort was made to observe typical mathematics instruction.

Classroom information. I collected class information of two school years, 2009 and 2010. The information such as multi-grade class/single-grade class, and ELD/English proficiency class was collected from the school district office and teachers.

Data Analysis

Description of Variables

Each research question has different dependent variable and independent variable. Research questions 1 and 4 have student mathematics achievement scores as dependent variable, while research questions 2 and 3 have teaching practice and teacher knowledge as dependent variable. Variable described in this section are of two levels. The former are student level variables. The latter are teacher level variables. Here, variables of classroom are considered as teacher level variable.

Student level variables. MATHEMATICS is student mathematics achievement scores from each DAP test. MATHEMATICS 1 is student achievement score on the first DAP of the school year, MATHEMATICS 2 is for

the second one, and MATHEMATICS 3 is for the third one. Since students from school year 2009 and 2010 took the same test set, the year is not considered in the analyses. That is, MATHEMATICS 1 in school year 2009 and 2010 is coded MATHEMATICS 1 without considering the year difference. MATHEMATICS 1, 2, and 3 has missing values in both years. The Hierarchical Linear Model (HLM) software allows for missing observations at first level (Raudenbush, Bryk, Cheong, Congdon, & Toit, 2004). MATHEMATICS is used as dependent variable for research question 1 and 4. The means of MATHEMATICS along with their standard deviations are provided in Table 4.

Table 4

Means and Standard Deviations for MATHEMATICS

	School Year 2009		School Year 2010		Total
	N (n= 127)	Mean (SD)	N (n=172)	Mean (SD)	
MATHEMATICS 1	119	42.73 (12.15)	172	43.62 (12.70)	291 (n=299)
MATHEMATICS 2	124	47.09 (14.72)	159	48.80 (14.70)	283
MATHEMATICS 3	127	53.26 (16.98)	155	57.74 (15.71)	282

Student English proficiency level (ELL) is a dummy coded variable with ELL coded as one and non-ELL as zero. As noted earlier, almost half of participant students are ELLs. Table 1 describes the number of ELLs in fourth graders. ELL used as a predictor in level-2 for research question 4.

Teacher level variables. Teaching practice (PRACTICE) is a continuous variable measured by CLASS instrument. The number of observation using CLASS instrument varied from one to four times in 2009-2010 and from four to eight times in 2010-2011. Also, the interval of observation was different for each teacher. However, these issues are considered in the analyses because HLM allows missing time point and different intervals of measures among participants (Sloane, Helding, & Kelly, 2008). Variable, PRACTICE was used for research question 2, 3, and 4. The school year, the type and the category of CLASS scores used for PRACTICE, however, is different according to the research questions. Table 5 describes CLASS scores used for each question. Means and standard deviations of CLASS scores are provided in Table 6.

Table 5

Description of CLASS scores Used for Research Questions

Research question	School Year	Type of Data	Category of CLASS
2. Change of Teaching Practice	2010	Raw score	Total CLASS measure, and three domain measures
3. Relationship between Teaching Practice and Knowledge for Teaching Mathematics	2009, 2010	Mean	Total CLASS measure, and three domain measures
4. Effect of Teacher Knowledge and Teaching on Student Mathematics Achievement Growth	2009, 2010	Mean	Total CLASS measure

Table 6

Means and Standard Deviations for the CLASS scores

	School year 2009		School year 2010		Total	
	M	SD	M	SD	M	SD
CLASS	32.70	10.88	34.96	7.90	33.69	9.61
Sub-Domain						
Emotional Support	12.13	3.35	13.79	3.08	12.86	3.29
Classroom Organization	12.31	3.94	12.63	2.94	12.45	3.49
Instructional Support	8.25	4.83	8.54	3.01	8.38	4.08

Knowledge for teaching mathematics (TKNOWLEDGE) is teachers' test score from a test with LMT, DMI and algebraic items. TKNOWLEDGE is used for research question 3 and 4 as a criterion and predictor. Means and standard deviations of TKNOWLEDGE are provided in Table 7.

Table 7

Means and Standard Deviations for the TKNOWLEDGE

	School year 2009		School year 2010		Total	
	M	SD	M	SD	M	SD
TKNOWLEDGE	23.01	5.24	29.29	6.11	25.78	6.37

COMBO is a dummy coded variable with multi-grade classroom coded as one and single-grade classroom as zero. The number of multi-grade classroom and single-grade classroom can be found in Table 2. The variable, COMBO is used for research question 2 as a predictor.

ELL_CLASS is a dummy coded variable. The basic and intermediate classes which have ELLs are coded as one and English proficiency classes are coded as zero. Class English proficiency level in school year 2010 was used. Eight are non-ELL classes, and 6 are ELL classes among fourteen classes in school year 2010. The variable, ELL_CLASS is used in research question 2 as a predictor.

Statistical Procedure

Each research question has a different statistical model in the analyses. For the first and fourth question which is about student achievement growth, I

employed three-level hierarchical linear model which allows one to analyze repeated measures nested within each student and students nested within each teacher. For the second question about teaching practice with repeated measures nested within each teacher, I used a two level hierarchical linear model. For the third question about the relationship between teaching and knowledge for teaching mathematics, a linear regression analysis and multiple regression analysis are employed. I provide more information on the models below. I used SPSS 19 to clean data and for the regression analyses and HLM 6 to estimate two and three level hierarchical linear models.

In research question 1, I examined the growth rate of student achievement considering the characteristic of the data such as repeated measures and students nested within class. For this analysis, I used Hierarchical Linear Model (Raudenbush, & Bryk, 2002). Since I focus on the students' achievement growth over time without paying much attention to student background or teacher variables in this question, I used unconditional model for level 2 and level 3. The unconditional model can “provides important statistics for studying individual growth, including the partitioning of variability in the individual growth parameters into level 2 and level 3 component” (Raudenbush, & Bryk, 2002, p. 241). The model is followed;

At Level 1,

$$Y_{tij} = \pi_{0ij} + \pi_{1ij} (TIME_{tij}) + e_{tij} \quad (3.1)$$

Where the Y_{tij} is MATHEMATICS at time t for student i in classroom j . The $TIME_{tij}$ is the amount of time in months elapsed from the first data DAP, which is

centered at the month of the first DAP test to have a meaningful intercept. While, π_{0ij} represents student i's initial mathematics score on DAP and π_{1ij} represents student i's growth rate across each DAP. The e_{tij} captures random error in the mathematics achievement scores around student i's growth trajectory, where $e_{tij} \sim N(0, \sigma^2)$.

At level 2,

$$\pi_{0ij} = \beta_{00j} + r_{0ij} \quad (3.2)$$

$$\pi_{1ij} = \beta_{10j} + r_{1ij} \quad (3.3)$$

Where the β_{00j} captures the average initial mathematics score within classroom j, and β_{10j} is the average growth rate within classroom j. The r_{0ij} and r_{1ij} are random effects that capture the individual deviation from the average.

And at Level-3,

$$\beta_{00j} = \gamma_{000} + u_{00j} \quad (3.4)$$

$$\beta_{10j} = \gamma_{100} + u_{10j} \quad (3.5)$$

Where the γ_{000} is the overall average initial mathematics score; and γ_{100} is the overall average monthly growth rate. The u_{00j} and u_{10j} are deviations that allow the average initial status and growth rates to vary across classrooms

In this analysis, I focus not only on the coefficients such as initial mathematics score on DAP, and average mathematics growth rate, but also on the covariance between the initial mathematics score and the growth rate. This covariance can tell if the data are fanning out or fanning in. Put differently, it shows if students who have higher initial achievement grow more or less over time. In addition, I investigated the proportion of variance within individual

student, among individual within classroom, and among classrooms (Raudenbush, & Bryk, 2002).

In research question 2, I investigated the change of teaching with three analyses. First, the outcome is teaching practice (PRACTICE) repeatedly measured by CLASS instrument and the predictor is only time, which implies unconditional model in level 2. Second, level 1 is the same with the previous model but it is conditional model in level 2 which has COMBO and ELL_CLASS as predictors. Third, the general model is the same with the second analysis but the outcome is the sub-domains of CLASS measures which are emotional support, classroom organization and instructional support. These sub-domains are examined separately with level-2 predictors, COMBO and ELL_CLASS.

The model for the first analysis is as follow:

At Level-1:

$$Y_{ti} = \pi_{0i} + \pi_{1i} (\text{TIME}_{ti}) + e_{ti} \quad (3.6)$$

Where the Y_{ti} is PRACTICE at time t of teacher i , and the TIME_{ti} , is the amount of time in weeks elapsed from the beginning of school year. While π_{0i} and π_{1i} represent teacher i 's initial status and weekly growth rate across a school year, respectively. The e_{ti} captures the deviation between teacher i 's score at time t and their linear growth trajectory, where $e_{ti} \sim N(0, \sigma^2)$.

And at Level-2,

$$\pi_{0i} = \beta_{00} + r_{0i} \quad (3.7)$$

$$\pi_{1i} = \beta_{10} + r_{1i} \quad (3.8)$$

Where, β_{00} and β_{10} represent the average initial status and growth rate of PRACTICE, respectively. While, r_{0i} captures deviations between teacher i 's initial status and the average initial PRACTICE score, and r_{1i} captures deviations between teacher i 's growth rate and the average growth rate. This model allows us to understand the intercept and growth, and the variability in them as well.

The model for the second analysis has two predictors, COMBO and ELL_CLASS in level-2. Level 1 is the same as the first analysis, which is the equation (3.6). The Level-2 model is presented.

At Level-2,

$$\pi_{0i} = \beta_{00} + \beta_{01} (\text{COMBO}) + \beta_{02} (\text{ELL_CLASS}) + r_{0i} \quad (3.9)$$

$$\pi_{1i} = \beta_{10} + \beta_{11} (\text{COMBO}) + \beta_{12} (\text{ELL_CLASS}) + r_{1i} \quad (3.10)$$

The result of this analysis can tell if the multi-grade class and English level of class influence teachers' initial teaching practice, and the change of it. Since COMBO and ELL_CLASS are dummy coded variables, the coefficient β_{01} and β_{02} are the COMBO gap effect and the ELL_CLASS gap effect on initial CLASS, respectively, while β_{11} and β_{12} are the COMBO gap effect and the ELL_CLASS gap effect on growth rate, respectively.

The third analysis focuses on the effect of COMBO and ELL_CLASS on the emotional support, classroom organization and instructional support separately. Hence, the level-2 model is the same with the second analysis, which is the equation (3.9) and (3.10). Also, level-1 is the same with the previous two models, which is equation (3.6), however, the outcome Y_{ti} is not total CLASS scores. Y_{ti} is CLASS sub-domains. In this analysis, I examined if emotional

support, classroom organization and instructional support separately changed during a school year.

The result of research question 2 allows us to tell the change of teaching practice in detail if the overall teaching practice changed, the emotional support changed, or the instructional support changed over time with the effect of grade combination and English level of the class.

In research question 3, a linear regression was employed to evaluate the prediction of teachers' mathematics knowledge for teaching on their actual teaching. For this, the predictor is teachers' knowledge test score (TKNOWLEDGE) and the criterion is teaching practice score (PRACTICE). Also, I evaluated the prediction of teaching practice (PRACTICE) on teacher knowledge (TKNOWLEDGE) using multiple regression analysis. Since CLASS sub-domains (emotional support, classroom organization, and instructional support) are predictors, multiple regression is employed. The zero order correlation and partial correlation were examined to evaluate the effect of each sub-domain on teacher knowledge controlling for other sub-domains.

Research question 4 is the conditional model from question 1 to investigate the effect of mathematics knowledge for teaching (TKNOWLEDGE) and their teaching practice (PRACTICE) on student mathematics growth. Hence, the level-1 model is remained the same with research question 1 (equation 3.1). Level 2 have individual student variable, ELL because almost half of the students in this school district are ELLs. Level 3 have teacher level variables such as COMBO and ELL_CLASS. Level-2 and 3 models are presented below.

At Level-2,

$$\pi_{0ij} = \beta_{00j} + \beta_{01j} (\text{ELL}_{ij}) + r_{0ij} \quad (3.11)$$

$$\pi_{1ij} = \beta_{10j} + \beta_{11j} (\text{ELL}_{ij}) + r_{1ij} \quad (3.12)$$

β_{00j} : The average initial mathematics score of non-ELLs in class j;

β_{01j} : The ELL gap effect on the average initial mathematics score within class j;

r_{0ij} : The individual deviation of initial score from the average initial mathematics score of class j;

β_{10j} : The average mathematics growth rate of non-ELLs in class j;

β_{11j} : The ELL gap effect on the average growth rate within class j; and

r_{1ij} : The individual deviation of growth rate from the average growth rate of class j.

And at Level-3,

$$\beta_{00j} = \gamma_{000} + u_{00j} \quad (3.13)$$

$$\beta_{01j} = \gamma_{010} \quad (3.14)$$

$$\beta_{10j} = \gamma_{100} + \gamma_{101} (\text{TKNOWLEDGE}_j) + \gamma_{102} (\text{PRACTICE}_j) \quad (3.15)$$

$$\beta_{11j} = \gamma_{110} \quad (3.16)$$

γ_{000} : The average initial mathematics score of non-ELLs across all classes (grand mean of the initial mathematics score) having a teacher with TKNOWLEDGE=0 and PRACTICE=0;

u_{00j} : The deviations that allow the average initial mathematics score to vary across classes;

γ_{010} : The ELL gap effect on the initial mathematics score across all classes;

γ_{100} : The average mathematics growth rate of non-ELLs across all classes having a teacher with TKNOWLEDGE = 0 and PRACTICE=0.

γ_{101} : The impact of teacher knowledge on the average mathematics growth rate of non-ELLs across all classes;

γ_{102} : The impact of teaching practice on the average mathematics growth rate of non-ELLs across all classes; and

γ_{110} : The ELL gap effect on the average mathematics growth rate across all classes.

In the research question 4, the coefficients are examined to evaluate the influence of the predictors. The variability of student's initial status within a classroom and between classrooms was examined as well.

Summary

This chapter described the sample, data collection, procedure, variables used in the study. The chapter also described the analytical models to assess change in teaching practice, and the impact of teacher level factors on student achievement growth. In chapter five, I discuss the results of the empirical investigation.

Chapter 4

RESULTS

The main object of this study is to evaluate students' mathematics achievement growth and the effect of teacher knowledge and teaching on it. Here I focus on 4th grade students' mathematics learning. To evaluate teacher influence on student learning, teaching practice and its change was examined. Also, the relationship between mathematics teaching and teacher's knowledge for teaching mathematics was examined. This chapter provides a graphical depiction of classroom learning in mathematics and describes the results of analyses for the four research questions separately.

Examining Student Growth Graphically

In this section I examine a graphical rendering of student mathematics achievement and their growth. I first do this by examining the intercepts and slopes of all participating students. Then I look at this achievement and growth classroom by classroom.

In Figure 3, the graph displays enormous within grade level variability in mathematics achievement. This graph mainly considers the effect of ELL factor on student monthly mathematics achievement growth. The effects of knowledge for teaching mathematics and teaching practice were not included in this graph. To begin, average student achievement at the beginning of fourth grade favors non-ELL students by nearly 12 points. In addition, non-ELL students grow slightly more than three times as fast as their ELL peers. Some, but very few, ELL students begin above their non-ELL peers (those students designated in red).

The range in entry level achievement is staggering at more than 60 points. Finally, at the current estimated rates of growth ELL students would require more than 26 months with the present teaching practice to reach the average entry level achievement of their non-ELL peers.

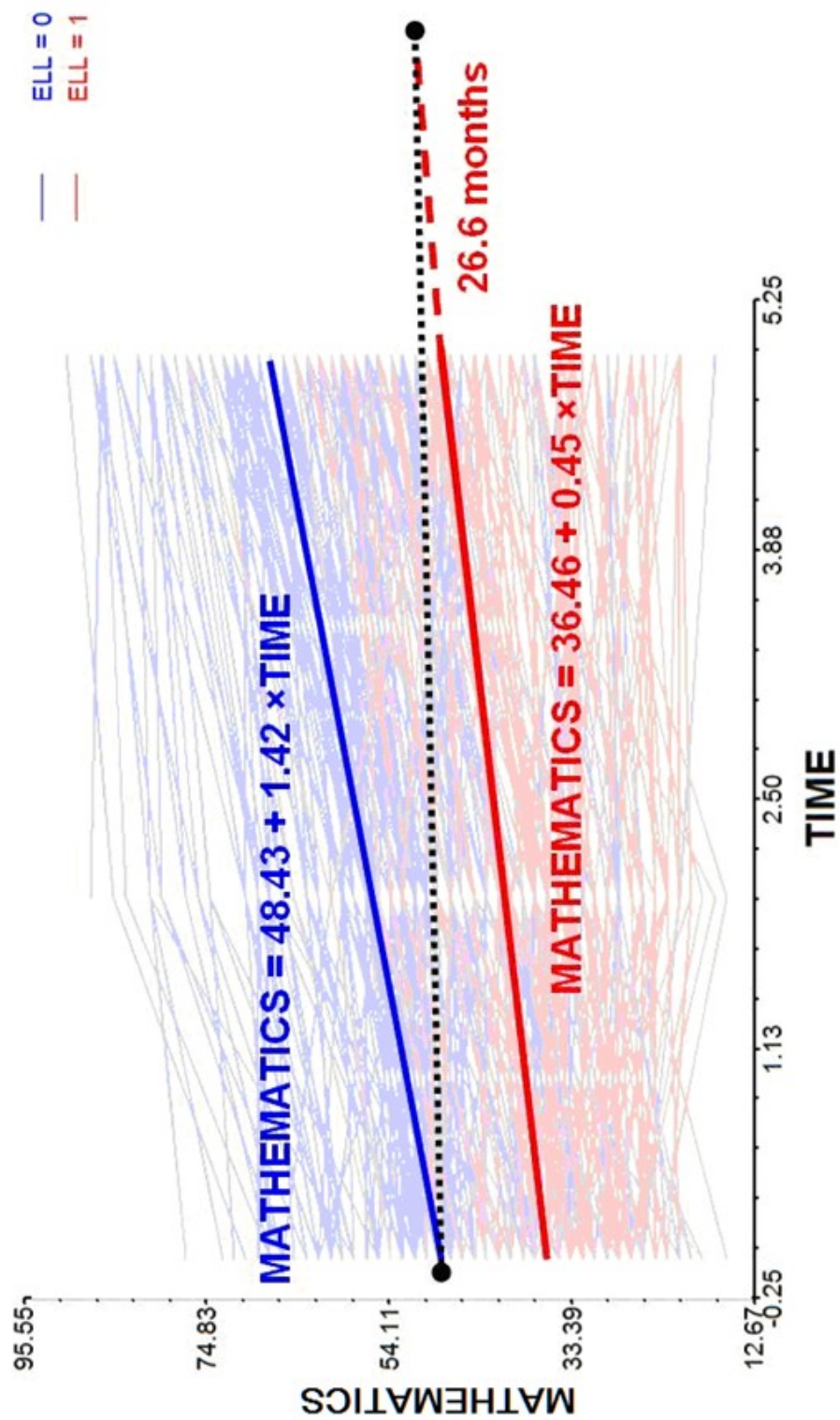
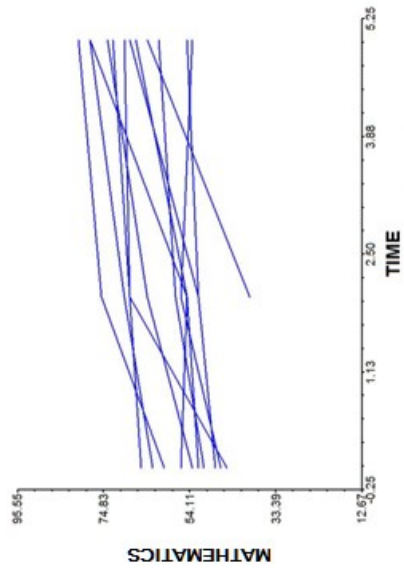


Figure 3. Student Mathematics Learning for All Participants

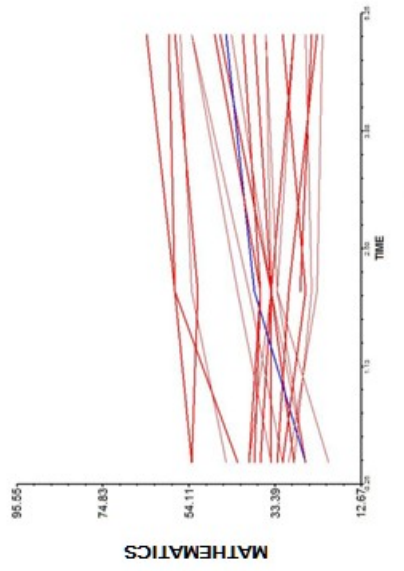
Classroom by Classroom Analysis

The sixteen participating classrooms are labeled 1 through 16 (level-id 1-16) (Figure 4). When I look at each classroom individually I note that some classrooms appear quite small in number (e.g., classroom 13). In reality this is not the case. What this actually indicates, in the case of classroom 13, is that of the more than 35 students in this classroom only 5 were fourth graders. Moreover, all of these fourth graders were second language learners.

Lev-id 1



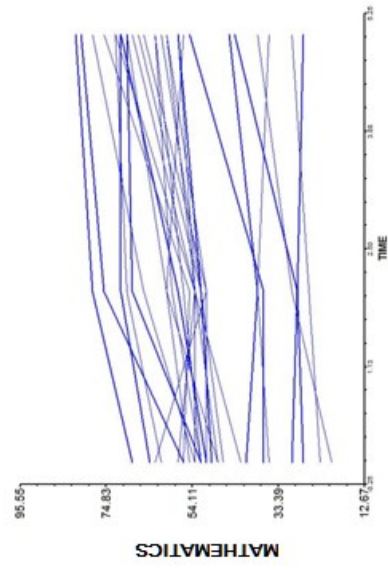
Lev-id 2



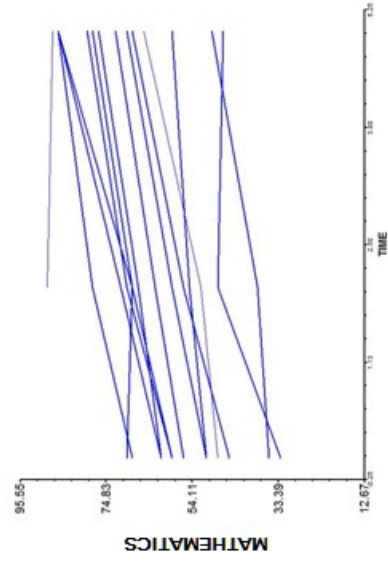
ELL = 0
ELL = 1



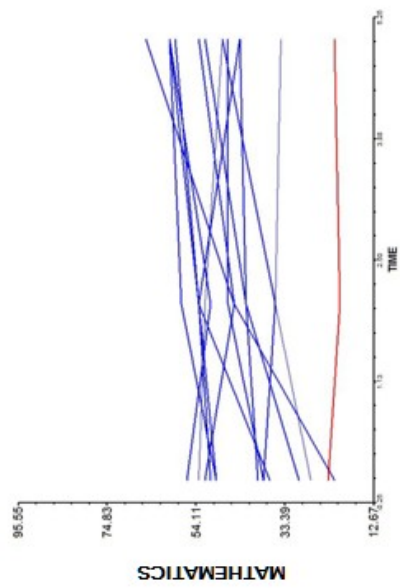
Lev-id 3



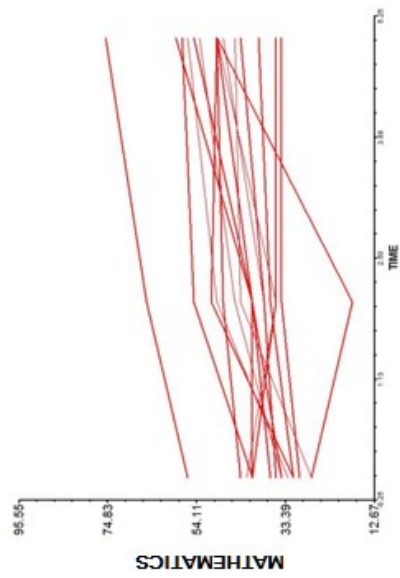
Lev-id 4



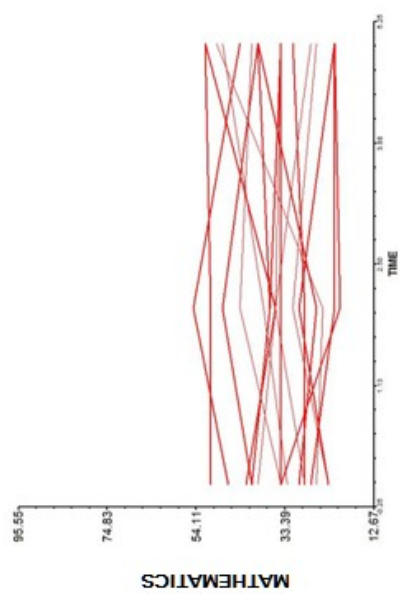
Lev-id 5



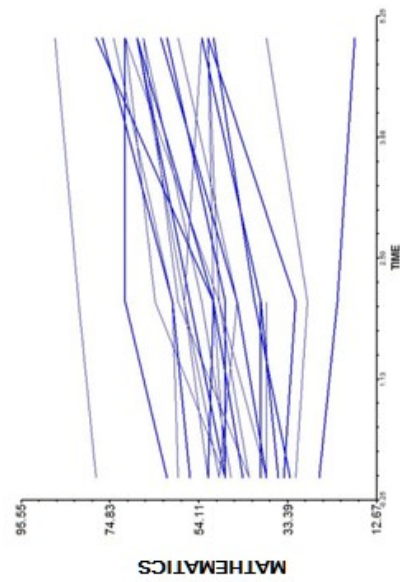
Lev-id 6



Lev-id 7

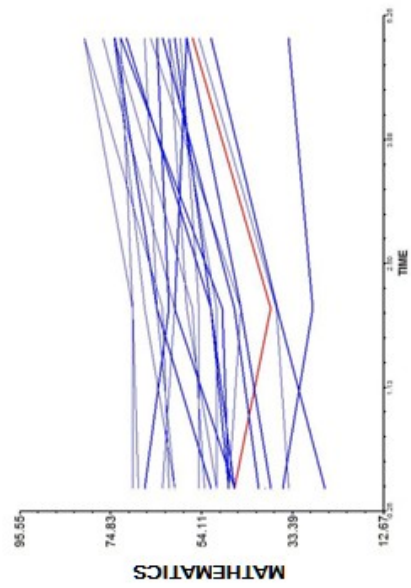


Lev-id 8

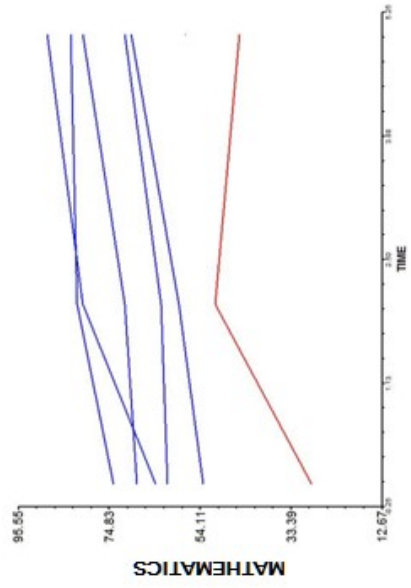


ELL = 0
ELL = 1

Lev-id 9

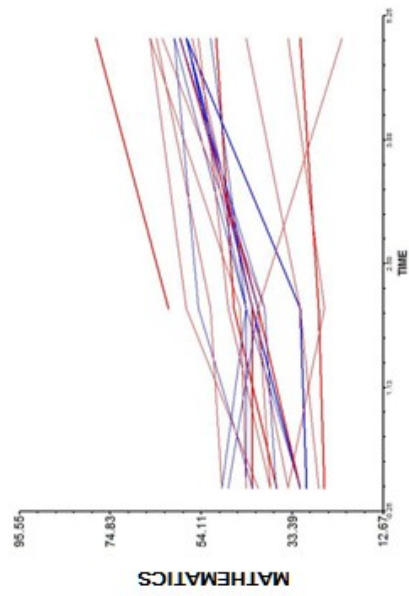


Lev-id 10

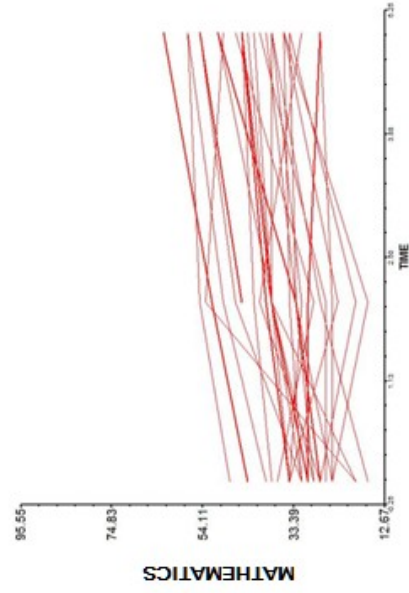


ELL = 0
ELL = 1

Lev-id 11



Lev-id 12



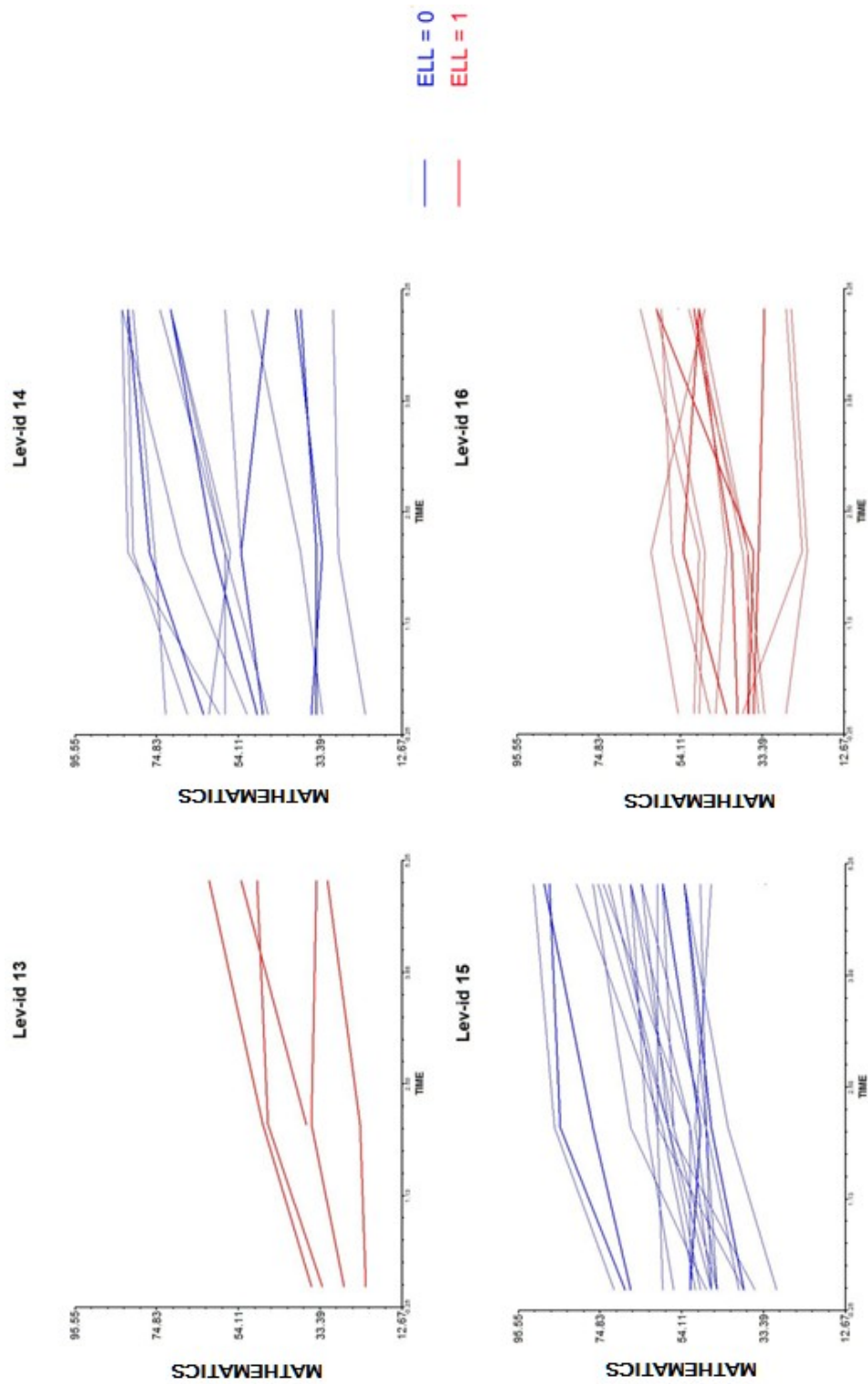


Figure 4. Student Mathematics Learning for Each Classroom

The next thing I notice is that the entry level of students is quite varied. Classes of non-ELL students demonstrate higher average achievement at the first testing date in the fall of fourth grade. However, there is considerable variation in entry level test score within classrooms (e.g., two students in classroom 3 score below 20 points while two others score above 60 points). This degree of variability makes the tailoring of instruction to fit the individual learning needs of students incredibly difficult. Finally, I also note that the rates at which student develop mathematically also varies considerably within classrooms irrespective of the first language of students. Classroom by classroom, non-ELL students enter with higher average measured achievement and grow at faster rates. I now examine the variability in these entry level scores and the variability in growth rates and their relationship to the research questions.

Student Mathematics Achievement Growth

The first research question is how students' mathematics achievement has changed for a school year. For this question, I look at fourth grade mathematics growth with unconditional three-level model. I begin at level-1 with an individual growth model of mathematics score on DAP test at time t for student i in classroom j :

At Level-1,

$$Y_{tij} = \pi_{0ij} + \pi_{1ij} (\text{TIME}_{tij}) + e_{tij} \quad (4.1)$$

Where, Y_{tij} is the outcome, MATHEMATICS, at time t for student i in classroom j . The time variable (TIME_{tij}) is defined as the amount of time in months that had elapsed from the first data-collection time. Because TIME_{tij} is

centered at the beginning of DAP test, it is 0 at first DAP (2 months after school starts), 2 at second DAP (4 months after school starts), and 5 at third DAP (7 months after school starts omitting winter break). The gap between test dates is accounted for in the analyses. The π_{0ij} is the initial mathematics score of student_{ij}, that is, the expected outcome for the student_{ij} at first DAP test; and π_{1ij} is the growth rate for student_{ij}.

At level 2,

$$\pi_{0ij} = \beta_{00j} + r_{0ij} \quad (4.2)$$

$$\pi_{1ij} = \beta_{10j} + r_{1ij} \quad (4.3)$$

And at Level-3,

$$\beta_{00j} = \gamma_{000} + u_{00j} \quad (4.4)$$

$$\beta_{10j} = \gamma_{100} + u_{10j} \quad (4.5)$$

β_{00j} represents the average initial mathematics score within classroom j, while γ_{000} is the overall average initial mathematics score; β_{10j} is the average growth rate within classroom j, while γ_{100} is the overall average growth rate.

The results in Table 8 show a strong positive growth averaged across all students and classrooms. The estimated initial mathematics score is 43.31. The average monthly growth rate is estimated at 2.55 points on a scale from 0 to 100. Table 9 describes the decomposition of the variance in the student initial mathematics score and growth into their within and between classroom components. The χ^2 statistics for these variance components indicate significant variation among students within classrooms for initial mathematics score ($\chi^2 = 1473.68$, $df = 267$, $p < 0.0001$) and their monthly growth rate ($\chi^2 = 397.06$, $df =$

267, $p < 0.0001$) as measured by the DAP test. There is also significant variation in the classroom means for the initial mathematics score ($\chi^2 = 167.69$, $df = 15$, $p < 0.0001$) and the monthly classroom mean growth rate ($\chi^2 = 69.23$, $df = 15$, $p < 0.0001$).

Table 8

Fixed Effect at Three-Level Analysis of Mathematics Score on DAP Test with Unconditional Model

Fixed Effect	Coefficient	se	t-ratio
Average initial mathematics score	43.31*	2.04	21.24
Average monthly growth rate	2.55*	0.21	11.93

* $p < 0.0001$

Table 9

Random Effect at Three-Level Analysis of Mathematics Score on DAP Test with Unconditional Model

Random Effect	Variance Component	df	χ^2	p-value
Level 1				
Temporal variation	24.17			
Level 2 (students within classroom)				
Individual initial mathematics score	87.29	267	1473.68	0.000
Individual growth rate	0.92	267	397.06	0.000
Level 3 (between classrooms)				
Classroom mean score	59.55	15	167.69	0.000
Classroom mean growth rate	0.55	15	69.23	0.000

Based on the variance component estimates, I computed the percentage of variation that lies within and between classrooms for both initial mathematics score and monthly growth rate (Table 10). About 41% of the variance in initial mathematics score lies between classrooms and 37% of the variance in growth rate lies between classrooms. Symmetrically, the within classroom estimates for initial status and growth rate are 59% and 63%, respectively.

Table 10

Variance Within and Between Classrooms at Three-Level Analysis of Mathematics Score on DAP Test with Unconditional Model

	Within Classrooms (%)	Between Classrooms (%)
Initial Dap score	59.45	40.55
Growth rate	62.59	37.42

As seen in Table 11 the estimated correlation between initial mathematics score and monthly growth rate is 0.35 for the students in the same classroom. This positive relationship implies that students who have high scores at first mathematics test tended to gain at a somewhat faster rate than their peers who begin the school year with lower performance. The relationship is similar and much stronger at the classroom level, 0.68.

For the mathematics score on the DAP measure, the estimated reliability for student initial status and growth rates are 0.82 and 0.32, respectively (Table 12). Reliability measures are 0.89 and 0.75 for the estimated classroom initial

DAP score and growth rate, respectively. This indicates the quality of these estimates to capture initial status and growth respectively.

Table 11

Variance-Covariance Components and Correlation among the Level-2 and Level-3 Random Effects at Three-Level Analysis of Mathematics Score on DAP Test with Unconditional Model

Variance-Covariance Component and Correlations		
Level 2	87.29	0.35
	3.11	0.92
Level 3	59.55	0.68
	3.92	0.55

Table 12

Reliability of Level-1 and Level-2 Random Coefficient at Three-Level Analysis of Mathematics Score on DAP Test with Unconditional Model

	Reliability
Level-1	
Individual initial mathematics score	0.82
Growth rate	0.32
Level-2	
Classroom mean of initial mathematics score	0.89
Classroom mean growth rate	0.75

Change of Teaching Practice

Second research question is if teachers' instructional practices (as measured by CLASS) change over the course of a school year. To investigate this question, I examined total CLASS score of 4th and 5th grade teachers. Also, I examined each domain of CLASS measures such as emotional support, classroom organization, and instructional support separately. First, using the overall CLASS score, I employed two-level HLM (unconditional model) to evaluate teacher's change as a function of time and the variability in initial status and growth over time, should such growth exist. Two-level model is presented.

At Level-1,

$$Y_{ti} = \pi_{0i} + \pi_{1i} (\text{TIME}_{ti}) + e_{ti} \quad (4.6)$$

And at Level-2,

$$\pi_{0i} = \beta_{00} + r_{0i} \quad (4.7)$$

$$\pi_{1i} = \beta_{10} + r_{1i} \quad (4.8)$$

Level-1 represents a growth model of PRACTICE. The Y_{ti} is measures of total CLASS score (PRACTICE) at time t of teacher i. The time variable, TIME_{ti} , was defined as the amount of time in weeks that had elapsed from the beginning of school year. The level-2 model represents the variability in the intercept (i.e., teacher status) and growth (i.e., teacher change). The estimated average intercept and weekly average growth rate for PRACTICE are 31.67 and 0.17, respectively (Table 13). Here, the average initial status of PRACTICE is significantly different from zero but the average weekly growth rate is not found to be

significantly different from zero. However, the latter approaches significance with an estimated p value of 0.07.

The estimates for the variances of individual initial status and growth parameters are 88.02 and 0.06, respectively (Table 14). The χ^2 statistics for initial status ($\chi^2 = 69.17$, $df = 13$, $p < 0.0001$) shows that I can reject the null hypothesis and conclude that teachers' PRACTICE varies significantly at the beginning of the school year. In parallel the χ^2 statistics for growth rate ($\chi^2 = 37.84$, $df = 13$, $p < 0.0001$) indicates significant variation in the estimated growth rates.

Table 13

Fixed Effect at Two-Level Analysis of PRACTICE with Unconditional Model

Fixed Effect	Coefficient	se	t-ratio
Average initial status	31.67*	2.80	11.30
Average growth rate	0.17	0.08	1.95

* $p < 0.0001$

Table 14

Random effect at Two-Level Analysis of PRACTICE with Unconditional Model

Random Effect	Variance Component	df	χ^2	p-value
Initial status	88.02	13	69.17	0.000
Growth rate	0.06	13	37.84	0.000
Level-1 error	24.40			

The Intraclass Correlation Coefficient (ICC) is found to be $88.02 / (24.40 + 88.02) = 78.30$. The ICC allows us to appropriately allocate the variance in

PRACTICE into their within- and between-person components. I find that a little over 22% of the variance resides the within-teachers (over time) and that the other 78% is allocated to the between teacher model.

For PRACTICE, reliabilities for initial status and growth rates are 0.80 and 0.64, respectively (Table 15). These estimates of reliability indicate the quality to the intercept and slope parameters as measures of initial status and growth. These estimates are high and well within the range of those found in the extant research literature. As one would expect reliability increases with increased measurement. In this case some teachers were observed on as many as eight occasions and most were observed more than four times with the average being somewhere between five and six.

Table 15

Reliability of Coefficient Estimate at Two-Level Analysis of PRACTICE with Unconditional Model

	Reliability
Initial status	0.80
Growth rate	0.64

Next, I consider the conditional model at Level-2 which estimates predictors of the variability in the intercept and the growth parameters of PRACTICE. I use two variables as predictors of this variability. The first is labeled COMBO. The second predictor variable is the ELL status of the classroom (ELL_CLASS).

Four models are presented. Each represents the study of the variability in teacher practice over the course of a school year. The first model (at level-1) is the same as the one presented above (equation 4.6) and represents a study of the variability in the overall CLASS score. The other three models capture the same relationships for each of the three subcomponents of the overall CLASS measure: Emotional Support, Classroom Organization, and Instructional Support. The general Level-2 model is presented symbolically below:

$$\pi_{0i} = \beta_{00} + \beta_{01} (\text{COMBO}) + \beta_{02} (\text{ELL_CLASS}) + r_{0i} \quad (4.9)$$

$$\pi_{1i} = \beta_{10} + \beta_{11} (\text{COMBO}) + \beta_{12} (\text{ELL_CLASS}) + r_{1i} \quad (4.10)$$

I set COMBO=1, if combination class, 0 if non-combination class, and ELL_CLASS = 1, if ELD classroom, 0 if proficient classroom. Thus, β_{00} and β_{10} represent the average intercept and slope for non-combination classroom and English proficiency classroom (COMBO = 0, ELL_CLASS=0), respectively.

As can be seen in Table 16 the intercept estimates of initial status for models 1, 2, 3, and 4 are significantly different from zero. These results parallel the one presented earlier for the unconditional model with overall CLASS score. The slope (or growth) estimate is found to be significant for the Emotional Support model. Teacher's weekly growth rate in a single-grade and English proficiency classroom is 0.09 for Emotional Support. On average, such teachers change their emotional support about 0.09 points by week. However, much like the overall model the estimated growth parameter is not found to be significantly different from zero in the case of the overall CLASS measure, Classroom Management and Instructional Support models. Results of the four conditional

level-2 models are presented below (Table 16 and 17). These models include the influence of the predictor variables noted above: multi-age, multi-grade classrooms (represented as COMBO) and the first language status of the classroom (represented by the variable ELL_CLASS).

Table 16 displays the estimated fixed effects for the four models. COMBO coefficients represent COMBO gap effect on initial status and growth rate. The COMBO gap on initial status (the difference of initial status between multi-grade classroom and single grade classroom) was significantly different from zero for model 2 (Emotional support) and model 4 (Instructional support). On average, teachers in multi-grade classrooms started with lower initial scores than those in single-grade classrooms by 4.68 points on a scale from 0 to 40 for Emotional Support and 4.87 points on a scale from 0 to 30 for Instructional Support. Teacher's initial emotional support score is 13.24 and 8.56 ($=13.24 - 4.68$) in a single-grade classroom and a multi-grade classroom, respectively. Teachers' initial instructional support score is 9.09 and 4.22 ($=9.09 - 4.87$) in a single-grade classroom and a multi-grade classroom, respectively. Other than that, neither COMBO nor ELL_CLASS influenced initial status for the other two models of initial status. Moreover, these two variables do not predict the variability in the change measure (i.e., the slope) of the overall CLASS scores and its component subscores.

Table 16

Fixed Effect Estimates at Two-Level Analysis of Teachers' overall CLASS and its Sub-domain Measures from Four Conditional Models

Parameter	Model 1 (overall CLASS)	Model 2 (Emotional Support)	Model 3 (Classroom Organization)	Model 4 (Instructional Support)
Model for initial status				
INTERCEPT	35.11 (3.71)**	13.24 (1.08)**	12.74 (1.65)**	9.09 (1.40)**
COMBO	-12.16 (5.70)	-4.68 (1.66)*	-2.52 (2.54)	-4.87 (2.14)*
ELL_CLASS	4.13 (5.69)	2.85 (1.66)	-0.60 (2.54)	0.96 (1.07)
Model for growth rate				
INTERCEPT	0.17 (0.13)	0.09 (0.04)*	0.06 (0.05)	0.02 (0.06)
COMBO	0.13 (0.20)	0.03 (0.06)	0.04 (0.07)	0.06 (0.09)
ELL_CLASS	-0.13 (0.20)	-0.09 (0.06)	-0.04 (0.07)	-0.01 (0.09)

Note. Standard errors are in parentheses.

* $p < 0.05$, ** $p < 0.0001$

Table 17 presents the estimated variances for the random effects in the four models. The results show that in the four models, there are significant variations in the teachers' initial status and growth rate of CLASS measures and its sub-domains such as Emotional support, Classroom organization, and Instructional support.

Table 17

Random Effect Estimates at Two-Level Analysis of Teachers' overall CLASS and its Sub-domain Measures from Four Conditional Models

Parameter	Model 1 (overall CLASS)	Model 2 (Emotional Support)	Model 3 (Classroom Organization)	Model 4 (Instructional Support)
Initial status	70.61**	4.60*	14.79**	9.00**
Growth rate	0.08**	0.01*	0.01*	0.02**
Level1 error	24.39	3.69	3.97	4.51

* $p < 0.05$, ** $p < 0.0001$

Relationship between Knowledge for Teaching Mathematics and Teaching Practice

Third research question is how well mathematics knowledge for teaching (measured by a test using LMT items, DMI items and relational thinking item) predicts teaching practice (measured by CLASS instrument). And how well does teaching practice predict teacher's mathematics knowledge for teaching? First, I examined how well mathematics knowledge for teaching predicts their actual teaching. To answer this question, a linear regression analysis was conducted. The criterion is overall CLASS measure and the predictor is a score from a test to measure mathematics knowledge for teaching. The results indicate that there is

no significant linear relationship between teachers' score from a test for knowledge and CLASS measure, $F(1, 30) = 1.52, p = 0.23$. The 95% confidence interval for the slope, -0.22 to 0.88 contains the value of zero and therefore mathematics knowledge for teaching is not significantly related to their teaching practice. The correlation between the test score for knowledge and CLASS measure was 0.22. Approximately 5% of the variance of CLASS measure was accounted for by the linear relationship with the knowledge score.

Second, I examined how well teaching practice predicts their mathematics knowledge for teaching. For this question, a multiple regression analysis was conducted. The criterion is teacher's score from a test for measuring mathematics knowledge for teaching, while the predictors are sub-domain of CLASS measures such as Emotional support, Classroom organization, and Instructional support. The linear combination of measures of CLASS sub-domain was not significantly related to the teacher knowledge score, $F(3, 28) = 0.55, p = 0.65$. The sample multiple correlation coefficient was 0.24, indicating that approximately 6% of the variance of teacher knowledge measures in the sample can be accounted for by the linear combination of the measures of CLASS sub-domains.

Table 18 presents the relative strength of the individual predictors. All the bivariate correlations between CLASS sub-domains and knowledge measure were not statistically significant. Also the partial correlation between CLASS sub-domains and knowledge measure was not significant.

Table 18

The Bivariate and Partial Correlations of the CLASS Sub-Domains and Teacher Knowledge Measure

Predictors	Correlation between each predictor and knowledge measure	Correlation between each predictor and knowledge measure controlling for all other predictors
Emotional support	0.17	-0.01
Classroom organization	0.23	0.14
Instructional Support	0.19	0.05

Effect of Teacher Knowledge and Teaching Practice on Student Mathematics

Achievement Growth

The final research question is if teachers' mathematics knowledge for teaching and their teaching practice significantly predict students' mathematics achievement growth. Here, I consider the conditional model from question 1, which allows the estimation of the separate effects of student variable (individual ELL status), and teacher variables (mathematics knowledge for teaching and teaching practice) on monthly mathematics growth rate. Three-level model is presented. Now I estimate the impact on the initial status and growth for the chosen predictor variables.

At Level-1,

$$Y_{tij} = \pi_{0ij} + \pi_{1ij} (\text{TIME}_{tij}) + e_{tij} \quad (4.11)$$

At Level-2,

$$\pi_{0ij} = \beta_{00j} + \beta_{01j} (\text{ELL}_{ij}) + r_{0ij} \quad (4.12)$$

$$\pi_{1ij} = \beta_{10j} + \beta_{11j} (\text{ELL}_{ij}) + r_{1ij} \quad (4.13)$$

And at Level-3,

$$\beta_{00j} = \gamma_{000} + u_{00j} \quad (4.14)$$

$$\beta_{01j} = \gamma_{010} \quad (4.15)$$

$$\beta_{10j} = \gamma_{100} + \gamma_{101} (\text{TKNOWLEDGE}_j) + \gamma_{102} (\text{TEACHING}_j) \quad (4.16)$$

$$\beta_{11j} = \gamma_{110} \quad (4.17)$$

Level-1 represents growth model of the mathematics achievement at time t of student i in classroom j. Level-2 model is the variability of the growth parameters among students within classrooms. The effects of student ELL status is presented here. I hypothesized that student ELL status is related to student initial mathematics score and monthly growth rate. Because ELL status is dummy variable, the corresponding regression coefficients are ELL-status gap effects. Put differently, β_{01j} represents the difference between ELL and non-ELL on student initial mathematics score in classroom j; and β_{11j} represents the difference of ELL status effect on student monthly mathematics growth rate in classroom j. The negative coefficients are anticipated because ELL is coded 1 and non-ELL 0. Moreover, it was hypothesized that ELL students would start lower in average. As noted earlier these students who start lower also progress or grew more slowly.

The level-3 model captures and analyses the variability among classrooms in the four β coefficients. I hypothesize that teacher's mathematics knowledge for teaching (TKNOWLEDGE) and CLASS scores to measure teaching practice (PRACTICE) predict classroom growth rate. I also hypothesize that the effect of

student ELL status on initial mathematics score and growth rate is considered a fixed effect across all classrooms. In this model, there is one random effect per classroom, which is u_{00j} .

The estimated fixed effects for this three level model are described in Table 19. For non-ELL student ($ELL=0$), the predicted mathematics score on first DAP is 48.43. At the initial data collect point (2 month after school starts), the score difference of ELL to non-ELL students is -11.97 points. ELL students start out 11.97 points behind their non-ELL peers, which is $48.43 - 11.97 = 36.46$. The effect size is 2.93. The initial status of ELL students is significantly different and lower than non-ELL students.

Table 19

Fixed Effect of Student ELL Status, Teacher Knowledge for Mathematics Teaching, and PRACTICE on Student Mathematics Learning

Fixed Effect	Coefficient	se	t-ratio
Model for initial status			
Model for mean status of non ELL student			
INTERCEPT	48.43**	1.57	30.86
Model for ELL difference on initial status			
INTERCEPT	-11.97**	2.17	-5.52
Model for monthly growth rate			
Model for growth rate of non ELL student			
INTERCEPT	1.42*	0.69	2.05
TKNOWLEDGE	-0.03	0.02	-1.34
TEACHING	0.07**	0.01	5.20
Model for ELL difference on growth rate			
INTERCEPT	-0.97**	0.21	-4.70

* $p < 0.05$, ** $p < 0.0001$

The predicted monthly growth rate for non-ELL student with a teacher whose knowledge score is zero and CLASS measure is zero is 1.42. On average, such students gain about 1.42 points per month. The PRACTICE coefficient represents monthly mathematics growth rate for non-ELL students as a function of teaching practice (PRACTICE). This estimate is 0.07. That is, a unit increase in PRACTICE predicts an average increase of mathematics growth of 0.07 point each month for non-ELL students. Teacher knowledge for teaching mathematics has a small negative effect on student learning growth for non-ELL students but it is not significantly different from zero.

For ELL students, monthly mathematics growth rate is 0.97 points lower than that for non-ELL students. The average monthly gain for ELL student is $1.42 - 0.97 = 0.45$ points and about 3.15 points over the time interval under this study. In contrast, non-ELL students grow about 9.94 points over the time interval under this study.

Table 20 presents estimated random effects and related χ^2 statistics from the three level decomposition. The χ^2 statistics for initial status of individual student within classroom ($\chi^2 = 3025.29$, $df = 282$, $p < 0.0001$) suggest that students' mathematics score in the same classroom vary significantly at the beginning of the school year. In parallel the χ^2 statistics for initial status of classroom mean ($\chi^2 = 56.98$, $df = 15$, $p < 0.0001$) indicates significant variation in classroom mean at the beginning.

Table 20

Variance Decomposition from a Three-Level Analysis of the Effects of Student ELL Status, Teacher Knowledge for Teaching, and CLASS Measure on Student Learning

Random Effect	Variance Component	df	χ^2	p-value
Level-1 variance				
Temporal variance	30.03			
Level-2 (student within classrooms)				
Individual initial status	102.90	282	3025.29	0.000
Level-3 (between classrooms)				
Classroom mean status,	17.88	15	56.98	0.000

Summary

Here I assessed the influence of teacher effects on student mathematics achievement growth. I close the chapter with a brief summary of each research question.

Research question 1: How students' mathematics achievement has changed for a school year?

There was a strong positive mathematics achievement growth averaged across all students and classrooms. Students' initial mathematics score and their monthly growth rate within classroom varied significantly. Classroom initial means for mathematics score and the monthly growth rate of classroom mean varied as well.

Research question 2: Do teachers' instructional practices (as measured by CLASS) change over the course of a school year?

Teachers' initial CLASS measure and its change varied significantly. But the average weekly growth rate of overall CLASS measure was not found to be significantly different from zero although p-value was closed to 0.07. The weekly growth rate of emotional support, however, was significantly different from zero. And teachers in multi-grade classrooms had lower emotional support and instructional support for their students at the beginning of the school year.

Research question 3: How well does mathematics knowledge for teaching (measured by a test using LMT items, DMI items and relational thinking item) predict teaching practice (measured by CLASS instrument)? And how well does teaching practice predict teacher's mathematics knowledge for teaching?

In this study, I could not find a significant linear relationship between teachers' knowledge score and CLASS measure and the linear combination of measures of CLASS sub-domain (emotional support, classroom organization, and instructional support) was not significantly related to the teacher knowledge score, either.

Research question 4: Do teachers' mathematics knowledge for teaching and their teaching practice significantly predict students' mathematics achievement growth?

Teaching practice significantly predicts non-ELL students' average monthly growth rate of mathematics. However, the effect of teachers' knowledge

for teaching mathematics on students' achievement growth is not statistically significant.

The following and the last chapter discuss the findings examined here and related them to the results from previous studies. In chapter 5, I also provide policy implications, some limitations, recommendations for future study, and overall conclusions.

Chapter 5

DISCUSSION AND CONCLUSION

A growing body of research shows that teachers effect on student mathematics achievement. Some researchers have examined the role of teacher knowledge and another has investigated teaching practice. However, relatively few such studies have focused on the effect of both teacher knowledge and teaching practice together on student achievement, or student achievement growth in mathematics. Moreover, many studies used proxy variables such as certification, degree or major for teacher knowledge and teachers' self-reported answers from surveys for measures of teaching practice. In terms of student achievement, these studies used student achievement score as a single time-point dependent variable and analyzed their data aggregated across entire classes or as disaggregated individual score. In addition, researchers analyzed teacher data and student data separately and drew conclusions from them.

Therefore, the primary purpose of this study was to investigate the effects of knowledge for teaching mathematics and teaching practice on student mathematics achievement growth. Knowledge for teaching mathematics was measured by a test and teaching practice was measured by observation. Student mathematics achievement was measured three times and multilevel analyses were employed to analyze the resulting hierarchical data. The following chapter is divided into five sections: 1) discussion with a summary of the findings, 2) their implications for policy, 3) study limitations, 4) recommendations for future study, and 5) conclusions.

Discussion

This section explains important findings from HLM and regression analyses. More details about and explanations of specific findings are discussed in terms of each research question.

Student Mathematics Achievement Growth

The average student initial mathematics score on DAP test is 43.31 points and the estimated monthly growth rate is 2.55 points; both are statistically and meaningfully significant as expected. Here, I would like to focus on the variability in the initial status and growth rate within and between classrooms. Student initial mathematics score and monthly growth rate significantly varied in the same classroom. Fifty nine percent of the variance in initial mathematics score and 63% of variance in growth rate resided within classroom. The variability within classroom comes from student individual level. In the study I have student English proficiency level as individual level variable to partially explain the variability within classroom in research question 4. In addition, classroom averaged initial status and monthly growth rate significantly varied. Forty one percent of variance in initial status and 37% of variance in growth rate is from between classrooms. To explain the variability that lies between classrooms, classroom level factors such as teacher knowledge, teaching practice, grade combination status and classroom English level were included and examined in research question 2, 3, and 4.

In the same classroom, student initial mathematics score and monthly growth rate was correlated (0.35) and such correlation is stronger at the classroom

level (0.68). That is, students with higher achievement scores at the beginning of the school year learned faster than those who had lower scores. This is more obvious at classroom level. The classes with higher initial mean on DAP test had faster mean growth. This implies that student learning gaps would get bigger and bigger during school year as the classes were grouped by the language status of the students. There is a clearly inequitable result associated with this grouping policy.

Change of Teaching Practice

Initial teaching practice was 31.67 and weekly growth rate was estimated at 0.17 (the growth rate is not significantly different from zero). The initial status of teaching practice significantly varied at the beginning of the school year and the estimated growth rate varied as well. About 80% of variance lay between teachers and 20% within teachers. The variability within teacher implies that the same teacher teaches differently each time (when measured by the CLASS instrument).

Teaching practice presents differences and similarities in a single-grade class and multi-grade class. In multi-grade classrooms, teachers have less emotional support in their teaching at the beginning of the school year. That is, when teachers have multi graders in the classroom, (1) they may have a lack of awareness of and responsiveness to students' emotional need, (2) they may present less sensitivity to facilitate students' ability to actively explore and learn, (3) they will likely have less interaction with students, and (4) classroom activities are less likely to emphasize students' interests, and motivations. Moreover, in

multi grades classroom, teachers have less instructional support initially. Put differently, teacher and students may have less discussion and activities to promote students' higher-order thinking, and they do not focus much on understanding but rather on rote instruction. Also, teachers provide less feedback that expands student learning and understanding, and less language stimulation. On the other hand, in the multi grade classroom teachers' classroom organization was not significantly different from that in a single grade classroom from the beginning of the year. That is, teachers' monitoring of students' behavior, behavior expectations, instructional time management, organizing activities and encouraging students' engagement in the activities were not significantly different between multi-grade and single-grade classrooms. In sum, although teachers in multi grade classrooms organize their classroom effectively as much as teachers do in single grade classrooms, they provide less support for their students' needs in terms of emotional and instructional aspects.

Teachers' overall teaching practice does not change over the school year. Such a result can be partly due to the transformation of the interval of observation. Teachers were observed at different time points so the interval was calculated by weeks, which is more than likely too short an amount of time for them to change their practice. The unit of time, one week, may be too short to notice and distinguish the change of overall teaching practice.

On the other hand, teachers' emotional support increased weekly. Emotional support evaluated positive and negative climate; teachers' sensitivity of student academic and emotional needs; and the emphasis on students' interest,

motivations and students' autonomy in the classroom. Such interactions between teacher and students and among students increased by the temporal unit (the week). This result reflects that teachers are stricter at the beginning of the school year in order to establish classroom routines and norms. As time goes on, the teacher and her students know each other better and teachers have a better sense of their students' emotional needs. So they can support students better in terms of the emotional aspect. Teachers' instructional support and classroom organization skill, however, were not increased. Emotional bond between a teacher and students may not guarantee critical academic interaction. Although teacher's awareness of students' interest and classroom work atmosphere increased, the instructional discussion and quality of feedback remained the same. Also, teachers' behavior management, and instructional routine and time management did not change.

Relationship between Knowledge for Teaching Mathematics and Teaching Practice

I could not find a statistically significant relationship between knowledge for teaching mathematics and teaching practice in this study. Also, the sub-categories of teaching practice such as emotional support, classroom organization and instructional support were not related to knowledge for teaching mathematics. The findings imply that high knowledge for teaching mathematics alone does not guarantee good teaching. Teachers' knowledge for teaching mathematics is a necessary condition for exemplary teaching but it is not a sufficient condition. In parallel, a teacher who has high teaching practice does not guarantee high teacher

knowledge for teaching mathematics. The findings imply that, like in many other situations, knowing (knowledge for teaching mathematics) is one thing and doing (teaching mathematics) is often quite another.

This result is not consistent with previous studies (e.g., Ball, 1991; Borko et al., 1992; Hill et al., 2008; Kahan et al., 2003; Stein et al., 1990; & Putnam et al., 1992) which say that teaching is highly related with and to teacher knowledge. The different result demonstrated here may generate from the different analytic lens. The previous studies were case studies which can describe the case in detail. The current study, however, quantified the degree of knowledge and teaching practice. The details in teacher knowledge and teaching practice could be lost in the quantifying process in this study and perhaps generated this inconsistent result relative to the body of prior work. Or the description process in case studies does not manifest itself and is not interested in the overall teaching practice seen in larger samples.

Another possibility of the nonsignificant relationship between knowledge and teaching is what I measure as knowledge for teaching mathematics is only a part of teacher knowledge. For example, the test items measured special content knowledge, knowledge for content and student, and knowledge of content and teaching but rarely focused on common content knowledge, knowledge of curriculum, or horizon content knowledge which Ball and her colleagues (Ball et al., 2008) argue as component of teacher knowledge. Or the number of items in the test was not enough to measure knowledge for teaching mathematics. In other words, there may be measurement issues associated with the incompleteness of

the instrument used for measuring knowledge for teaching mathematics in the study. Either it did not measure knowledge for teaching mathematics fully or the number of items was not enough to measure the aimed categories of teacher knowledge.

Alternatively, the nonsignificant findings may have resulted from my choice of measurement tool of teaching practice. The CLASS tool focuses not only on instructional quality but also the emotional relationship between teacher and students, and the classroom organization. Four out of ten dimensions focused on teachers' emotional support and three dimensions scored classroom organization. Hence teaching practice measured by the CLASS tool is combined score for overall teaching practice, which is quite different to what the previous studies focused on as the teaching practice. Previous studies focused more on mathematics instruction without paying much attention to classroom climate and teachers' behavior management.

When narrowing down teaching practice as instructional support excluding emotional support and classroom organization in CLASS measures, the measure is still not related to teacher knowledge. This result can be due to the perspective gap on teaching. How mathematics education researchers see teaching in the previous studies might be different from what the CLASS tool focused as teaching in instructional support component.

Overall what the current study measures for teacher knowledge and teaching practice emphasize different aspects with previous studies. With the

measures I could not find a significant relationship between teacher knowledge and teaching practice in my study.

Effect of Teacher Knowledge and Teaching Practice on Student Mathematics Achievement Growth

To investigate student mathematics achievement growth, ELL status was used as individual level predictor, and knowledge for teaching mathematics and teaching practice were used as teacher level predictors. As results from individual level indicate, ELLs had significantly lower achievement scores in the beginning of the school year. Also, ELL status is a very powerful factor in predicting student learning trajectories. ELLs started 11.97 points behind and learn 0.97 points lower than their non-ELL peers every month. That is, the gap between ELLs and non-ELLs is 18.76 points at the end of fourth grade with the original difference (11.97) and monthly growth gap for seven month (0.97×7) which is a whole school year without summer and winter break. If this pattern is kept on student mathematics learning, the gap between ELLs and non-ELLs would keep getting bigger and bigger on into the fifth grade, because the difference of their starting line in fifth grade would be more serious than it is in fourth grade, and there still would be the growth rate difference to consider between the two groups.

From teacher level variables, knowledge for teaching mathematics did not predict student mathematics achievement growth but teaching practice was significantly related student mathematics learning trajectory. The nonsignificant result of teacher knowledge is consistent (Eisenberg, 1977) and inconsistent (Hill et al., 2005; Mullen, 1996; Rowan et al., 1997) with the previous studies which

examined the relationship between teacher knowledge and student achievement. Also, it is partly consistent and partly inconsistent with Carpenter and his colleagues' study (1988) which stated that teachers' knowledge about problem type, problem difficulties and student strategies were not significantly related with student mathematics achievement but teachers' prediction of student success in problem solving was significantly correlated with student achievement. In addition, the finding is partly consistent with Baumert and his colleagues' work (2010) which says that teachers' pedagogical content knowledge powerfully predicts student achievement but their content knowledge does not.

The findings with different aspects to previous studies may have resulted from the repeated measures of student achievement. This study measured student achievement three times and investigate the growth (change) in achievement over time. Most of previous studies measured student achievement once and used it as a single-time point dependent variable. In other cases, researchers measured student achievement twice and used one measure as a covariate and the other as dependent variable, or they used the difference between two measures as dependent variable.

Willett (1988) explained that "the very notion of learning implies growth and change" (p.346). Said differently, the change or growth of in student achievement implies that students have learned mathematics (measured by tests) in some intervening period of time. Because a one-time measure does not acknowledge any pre-existing differences in student ability, we do not know how much they changed with a score measured once. Hence, it is difficult to say that

teachers' higher knowledge related to student learning with students' test score measured only once. In addition, pre-post test gives only minimal information about human growth (Willett, 1989). This study with repeated measures, however, acknowledged students' original differences and focused on their growth as a function of time and teacher factors. In this case, although a student final score is lower than another student, his or her growth can be greater than the others. This growth can detect the effect of teacher factors on student achievement.

Alternatively, the different aspects of results with previous studies and nonsignificant results could be related to the measurement tools as noted earlier. The selected problem set to measure teacher knowledge is not enough to capture knowledge for teaching mathematics. The test set may need to include more items for special content knowledge, knowledge of content and student, and knowledge of content and teaching, or include items about common content knowledge, knowledge at the mathematical horizon, and knowledge of curriculum for teacher knowledge from Ball's teacher knowledge category (Ball et al., 2008).

On the other hand, teaching practice was a significant predictor for student mathematics achievement growth. Despite previous study noting that teaching practice did not predict student mathematics achievement (Rowan et al., 1997), the results extend the findings of Cohen and Hill (2000), and Fennema et al. (1996). Although the previous studies used student achievement measured once, the current study analyzed student mathematics achievement growth with repeated measures. For the estimated growth teaching practice played an essential role as a predictive indicator.

The growth model analysis indicated that a unit increase in teaching practice measured by CLASS instrument predicts 0.07 point growth in mathematics achievement for non-ELL students each month. The mean of teaching practice score in the study is 33.69 and the standard deviation is 9.61. Then the teacher whose CLASS score is two standard deviations below the mean has 14.47 points ($= 33.69 - 9.61 \times 2$) as teaching practice. The difference of teaching practice score between the teacher and one in the average is 19.22 points ($= 33.69 - 14.47$). Then their student achievement growth difference in those two classrooms is 1.35 points ($= 19.22 \times 0.07$) every month and 9.45 points for seven months which is a school year without including summer and winter break. More seriously there would be 18.9 points ($= 9.45 \times 2$) achievement differences in the classrooms, where teaching practice is two standard deviation above and below than the mean, for a year.

In summary, the results highlight several important points. Initial student mathematics achievement and monthly growth rate varied within classroom and between classrooms. To explain the variability, individual level variable and teacher level variables were included. Student ELL status was one of the factors to explain within classroom variability. ELL students had significantly lower initial mathematics achievement and monthly growth rates than non-ELL students. In addition, to explain between classroom variability, knowledge for teaching mathematics and teaching practice were included as teacher level factors. First, the relationship between teacher level factors was examined. Overall teaching practice did not change weekly. Specifically, instructional quality and classroom

organization skill did not increase but teachers' emotional support for students increased significantly by week. In the multi-grade classrooms, teachers' support for student emotional and instructional needs was lower than that in a single-grade classroom initially. Moreover, knowledge for teaching mathematics and teaching practice was not significantly related. Finally, teaching practice significantly influenced on student mathematics learning but knowledge for teaching mathematics did not.

Policy Implications

The current study examined the effect of teacher knowledge and teaching practice on student mathematics achievement growth with multilevel analysis. The findings have several policy implications. First, ELL students need to be supported for mathematics learning. The findings indicate that ELL students start with lower mathematics achievement and learn more slowly than non-ELL students. Also the learning gap was huge. In this school district, ELD classes, where most ELLs are, have targeted special education in English; a 4-hour Language Arts block (reading, grammar, writing, and conversation), while classes with English proficiency students have a two hour reading block. The different course structures afford different amounts of time at a teacher's discretion that can make less time for teaching and learning mathematics for ELLs. Although the district requires a minimum one hour of mathematics class every day, the different curricula can make the mathematics learning gap between ELLs and non-ELLs bigger. Thus, the school system and policy makers need to purposefully enhance ELLs' English and their mathematics learning as well.

Second, to combine different grade levels in the same classroom, school principals and the school district need to be more careful. The findings indicated that teachers' support for students' emotional and instructional need is significantly lower in grade combination classes from the beginning of the year. Thinking that teaching practice is a strong factor in predicting student mathematics achievement growth, the poor teaching practice due to combination of grade could hurt student mathematics learning. When combining multi-grades in a single physical classroom is deemed necessary because of economic issues or curricular issues, teachers need to understand the differences and difficulties of multi-grade classes in terms of both the instructional and emotional aspects. Also, teachers should be well prepared to address those difficulties in their classrooms.

Third, professional development programs and in-service teacher programs should provide teachers with the opportunities to reconsider and enhance their instructional quality and classroom management. The findings suggest that the overall teaching practice powerfully predicts student mathematics achievement. The overall teaching practice includes instructional quality (e.g., encouraging student higher order thinking and giving quality feedback), classroom organization (e.g., monitoring student behavior and instructional time management) and emotional support (e.g., positive classroom climate, and emphasizing student motivation). Regarding the three domains, teachers improve their emotional support for students during a school year but they rarely change their instructional quality and classroom organization. This implies that although teachers get to know students' interests and have more interaction with their

students as the year progresses, the interaction may not be related to mathematics teaching and learning. Rather, the increased interaction and support can be for the non-mathematics content areas. To balance the three domains (instructional support, emotional support and classroom organization) in teaching, teachers need to be aware of their instructional quality and classroom organization skill along with their emotional support for their students. Since the emotional support is relatively easily improved, teachers should try to enhance their instructional quality and classroom organization skill intentionally. Hence, teachers' effort to change their instructional quality and classroom organization is essential. Therefore, the professional development program and in-service teacher program need to give teachers the opportunity to reflect on their instructional support and classroom organization. That opportunity should provide teachers with the chance to improve their teaching practice.

Study Limitations

There are a number of limitations in this quantitative study. First of all, this study included only student English level (ELL or non-ELL) as an individual level variable but there are other individual level variables such as SES, parents' education or hours study at home. For example, students study mathematics in the home alone or with their siblings or parents. This work effects their achievement on tests and their learning at school. Hence, students' study time at home influences their achievement growth. Also, according to parents' education level or economic status, students can have different educational attention and time available to them from their parents. These individual differences are the

factors that relate to variability in achievement within classrooms and can serve as covariate in the analysis but these factors were not included in current study.

Second, the language test for student English level is not about English for learning mathematics. Rather it is a test for everyday English. Some of the problems associated with language and achievement testing include confusion associated with homophones, where the same word can be interpreted differently in daily life and a mathematics context. For example, a student might receive the following prompt: “Which of the following numbers is odd? 2, 4, 6, 7, 24”. The word “odd” has a different meaning in mathematics and everyday English. The number 24 may be considered to be a correct response because here it looks odd due to it being a two digit number. In mathematics “odd” refers to the form $n = 2k + 1$, where k is an integer in the mathematics context (Middleton, Lamas-Flores, & Guerra, in press). Students must know the word “odd” they used in their English test is different from the meaning of “odd” in mathematics. The English test score in the current study may not explain how well students understand the language in and for mathematics.

Third, for question 1 and 4, the number of teachers who participated in the study is not sufficient. Sixteen teachers provided data for the analyses presented in question 1 and 4. Although it is larger than ten so multilevel modeling is attractive (Hox, & Maas, 2002), the power could be low and the possibility of type II error increases due to the small sample size in level three. To compensate for this defect, I did not include many teacher variables (combination of grades, class English level, or years of teaching experience) in level-3 for research

question 4. In addition, the effects of teacher knowledge and teaching practice on ELL gap in monthly growth rate across all classes (to investigate teacher effects on ELLs) were not included in the model although for non-ELLs that was investigated.

Fourth, this study selected items to measure teacher knowledge for teaching mathematics from two instruments (LMT and DMI). These items were found to be unidimensional in an exploratory factor analysis and to demonstrate reasonably high reliability using Cronbach's alpha (.83). However, in the process of selecting and combining items for my test, I might have missed the original intention or goal of the test developers. Because LMT is multiple-choice items and DMI open ended items, the combination of those test items gave me more information. However, it might hurt the validity of my measure, while reducing the overall reliability found for any of the complete measures independently.

Lastly, this quantitative methodology did not examine discourse in mathematics classes closely because teaching practice was quantified. Specifically it was not able to address such questions as: (1) how does teaching practice as manifested in within classroom dialogue influence student mathematics achievement growth, (2) how do teachers' discourse function differently in multi-grade class and in a single-grade class or (3) how does the same teaching practice work differently to ELL and non-ELL.

Recommendations for Future Study

The findings and limitations serve to contextualize recommendations for future study. The current study includes only student English level as individual

level variable. Other individual level variables such as hours to study at home, parents' education level, or SES need to be measured and controlled to investigate the effect of teacher and school factors on student achievement growth. However, we should carefully consider how to collect and measure the information. In the previous studies such as Trends in International Mathematics and Science Study (TIMSS), researchers asked relatively simple self-report questions (e.g., the highest education level of either parent, the number of books in the home, availability of computers and a student desk in the home) (Martin, 2005). Because these factors are correlated with SES, they can give us an indication of SES, but such items are limited in their value. What we find from self-reported survey items tend to be far removed from what actually is or occurs (Wenglinsky, 2002). Thus, we need to carefully measure and analyze student individual level variables with instruments that are sensitive to situation and have high levels of both validity and reliability.

Student English level for and in mathematics should also be measured. English level is a powerful factor to predict student mathematics achievement growth for ELLs. However, there can be some significant differences between students' daily English and English used in mathematics. Moreover, to enhance mathematics learning for English learners we need to examine the relationship between their everyday English and English in mathematics. Thus, future studies should develop the instrument to measure English in and for mathematics. The data collected with these improved instruments should help us better understand ELLs' learning in mathematics.

To have more power and reduce the type II error rate, future studies should draw on larger sample sizes at the cluster level (here classrooms). In addition, with an adequate sample size in level three, more classroom level variables can be included in the model. We found that teaching practice in a multi-grade class and in a single grade class was not the same but we do not know whether grade combination effects on student mathematics learning directly or not, since the factor was excluded in the final analysis. Moreover, with larger sample sizes, we can understand the effect of teacher factors on ELLs' achievement across all classes, the examination of which was excluded from this study.

This study was designed post-hoc as an addendum to an ongoing piece of research and as such suffered from a number of design weaknesses that made the study exploratory in nature in contrast to an inferential design. However, the research is important in its own right and adds significantly to the extant research literature. In summary those designing future studies need to carefully attend to a number of critical issues during the design phase:

1. The design should be multilevel in nature and support insight theoretically;
2. The design should focus on student learning in mathematics within a school year if teacher effects are to be estimated relative to curricular implementations;
3. Qualitative cases should be developed along with the quantitative multilevel study;

4. Nested power analyses should be conducted a-priori to allow for more robust estimation. The number of theoretical variables to be modeled should be considered when designing the sample to be drawn;
5. Appropriate policy tractable variables need to be included along with critical covariates;
6. The critical role of teachers needs to be elaborated theoretically and appropriate variables need to be developed and operationalized;
7. The important role of ELL students and their learning needs should be attended to in the design of the future studies along with the cultural settings that support such learning;
8. Variation across classrooms should be anticipated and this variation needs to be theoretically grounded as it relates to the learning of mathematics in ELL settings.
9. The local policy setting must also be well understood before the research design is put in place. Such understanding impacts significantly what can be asked and answered, with and, by high quality research.

In close, it is important that policy makers along with mathematics education researchers better understand the needs of ELL learners as learners of mathematics. It is all too easy to envision these needs as weaknesses in individual students. These weaknesses could be conceived incorrectly as a deficit, relative to their non-ELL peers. Rather, we need to see these learning needs for what they are: the explicit and implicit needs of students in classrooms not yet built to

support maximal, let alone, optimal learning on the part of second language learners. It is all too easy to blame the victims here and the reader is cautioned against such inherently incorrect and truly flawed inferences.

Conclusion

After conducting this research I have understood even more the importance of teaching practice. Teachers teach students, and the teacher is the most important single policy tractable variable that can be moderated to influence student learning (Wright et al., 1997). From the teacher variables, teaching practice powerfully predicted student mathematics achievement growth. Thus, to improve student mathematics achievement, teaching practice needs to be emphasized and improved.

Teaching practice is not fixed for a teacher. Teaching practice varied within a teacher and teachers' emotional support for their students increased during a school year. However, instructional quality and teachers' classroom organization are challenges for teachers in this sample and likely need to be improved nationally. Therefore, school systems and education policy makers need to provide teachers with the chance to reflect on their teaching and change it within themselves.

REFERENCES

- Aaronson, D., Barrow, L., & Sander, W. (2007). Teacher and student achievement in the Chicago public high schools. *Journal of Labor Economics*, 25, 95-135.
- Ball, D. L. (1990). Prospective elementary and secondary teachers' understanding of division. *Journal for Research in Mathematics Education*, 21, 132-144.
- Ball, D. L. (1991). Research on teaching mathematics: Making subject matter knowledge part of the equation. In J. Brophy (Ed.), *Advances in research on teaching: Vol. 2. Teacher's subject-matter knowledge* (pp. 1-48). Greenwich, CT: JAI Press.
- Ball, D. L., & Hill, H. (2008). *Mathematical knowledge for teaching (MKT) measures: Mathematics released items 2008*. Retrieved from http://sitemaker.umich.edu/lmt/files/LMT_sample_items.pdf
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59, 389-407.
- Baumert, J., Kunter, M., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S. Neubrand, M., & Tsai, Y-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47, 133-180.
- Bell, C. A., Wilson, S. M., Higgins, T., & McCoach, D. B. (2010). Measuring the effects of professional development on teacher knowledge: The case of Developing Mathematical Ideas. *Journal for Research in Mathematics Education*, 41, 479-512.
- Borko, H., Eisenhart, M., Brown, C. A., & Underhill, R. G. (1992). Learning to teach hard mathematics: Do novice teachers and their instructors give up too easily? *Journal for Research in Mathematics Education*, 23, 194-222.
- Carpenter, T. P., Fennema, E., & Franke, M. L. (1996). Cognitively guided instruction: A knowledge base for reform in primary mathematics instruction. *The Elementary School Journal*, 97, 3-20.
- Carpenter, T. P., Fennema, E., Franke, M. F., Levi, L., & Empson, S. B. (1999). *Children's mathematics: Cognitively guided instruction*. Portsmouth, NH: Heinemann.
- Carpenter, T. P., Fennema, E., Peterson, P. L., & Carey, D. A. (1988). Teachers' pedagogical content knowledge of students' problem solving in

- elementary arithmetic. *Journal for Research in Mathematics Education*, 5, 385-401.
- Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C-P, & Loef M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal*, 26, 499-531.
- Cobb, P., Wood, T., Yackel, E., & McNeal, B. (1992). Characteristics of classroom mathematics traditions: An interactional analysis. *American Educational Research Journal*, 29, 573-604.
- Cohen, D. K., & Hill, H. C. (2001). Instructional policy and classroom performance: The mathematics reform in California. *Teachers College Record*, 102, 294-343.
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *The Journal of Human Resources*, 41, 778-820.
- Eisenberg, T. A. (1977). Begle revisited: Teacher knowledge and student achievement in algebra. *Journal for Research in Mathematics Education*, 8, 216-222.
- Even, R. (1993). Subject-matter knowledge and pedagogical content knowledge: Prospective secondary teachers and the function concept. *Journal for Research in Mathematics Education*, 24, 94-116.
- Fennema, E., Carpenter, T. P., Franke, M. L., Levi, L., Jacobs, V. R., & Empson, S. B. (1996). A longitudinal study of learning to use children's thinking in mathematics instruction. *Journal for Research in Mathematics Education*, 27, 403-434.
- Franke, M. L., & Kazemi, E. (2001). Learning to teach mathematics: Focus on student thinking. *Theory into Practice*, 40, 102-109.
- Goldhaber, D. D., & Brewer, D. J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis*, 22, 129-145.
- Graeber, A. O., Tirosh, D., & Glover, R. (1989). Preservice teachers' misconceptions in solving verbal problems in multiplication and division. *Journal for Research in Mathematics Education*, 20, 95-102.

- Hill, H. C., & Ball, D. L. (2004). Learning mathematics for teaching: Results from California's mathematics professional development institutes. *Journal for Research in Mathematics Education*, 35, 330-351.
- Hill, H. C., Ball, D. L., Blunk, M., Goffney, I. M., & Rowan, B. (2007). Validating the ecological assumption: The relationship of measure scores to classroom teaching and student learning. *Measurement: Interdisciplinary Research and Perspectives*, 5, 107-118.
- Hill, H. C. Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal of Research in Mathematics Education*, 39, 372-400.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 43, 371-406.
- Hill, H. C., Schilling, S. G., & Ball, D. L., (2004). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal*, 105, 11-30.
- Hiebert, J., & Wearne, D. (1993). Instructional tasks, classroom discourse, and students' learning in second-grade arithmetic. *American Educational Research Journal*, 30, 393-425.
- Hox, J.J. & Maas, C.J.M. (2002). Sample sizes for multilevel modeling. In: J. Blasius, J. Hox , E. de Leeuw & P. Schmidt (Eds.). *Social Science Methodology in the New Millennium*. Proceedings of the Fifth International Conference on Logic and Methodology.
- Kahan, J. A., Cooper, D. A., & Betha, K. A. (2003). The role of mathematics teachers' content knowledge in their teaching: A framework for research applied to a study of student teachers. *Journal of Mathematics Teacher Education*, 6, 223-252.
- Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal*, 27, 29-63.
- Lampert, M. (2001). *Teaching problems and the problems of teaching*. New Haven: Yale University Press.
- LaParo, K. M., Pianta, R. C., & Stuhlman, M. (2004). Findings from the prekindergarten year. *The Elementary School Journal*, 104, 409-426.

- Leinhardt, G., & Smith, D. A. (1985). Expertise in mathematics instruction: Subject matter knowledge. *Journal of Educational Psychology*, 77, 247-271.
- LoCasale-Crouch, J., Konold, T., Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2007). Observed classroom quality profiles in state-funded pre-kindergarten programs and associations with teacher, program, and classroom characteristics. *Early Childhood Research Quarterly*, 22, 3-17.
- Ma, L. (1999). *Knowing and teaching elementary mathematics*. Mahwah, NJ: Erlbaum.
- Martin, M. O. (Ed.). (2005). *TIMSS 2003 User guide for the international database*. Chestnut Hill, MA: Boston College.
- McCaffrey, D. F., Hamilton, L. S., Stecher, B. M., Klein, S. P., Bugliari, D., & Robyn, A. (2001). Interactions among instructional practices, curriculum, and student achievement: The case of standards-based high school mathematics. *Journal for Research in Mathematics Education*, 32, 493-517.
- McClain, K., & Cobb, P. (2001). An analysis of development of sociomathematical norms in one first-grade classroom. *Journal for Research in Mathematics Education*, 32, 236-266
- Middleton, J. A., Llamas-Flores, S., & Guerra, P. (in press). English language learning and learning academic language in mathematics. In B. Arias & C. Faltis (Eds.), *Research in Second Language Learning: Academic Language in Second Language Learning*. Charlotte, NC: Information Age Publishing.
- Monk, D. H. (1994). Subject area preparation of secondary mathematics and science teachers and student achievement. *Economics of Education Review*, 13, 125-145.
- Mullens, J. E., Murnane, R. T., & Willett, J. B. (1996). The contribution of training and subject matter knowledge to teaching effectiveness: A multilevel analysis of longitudinal evidence from Belize. *Comparative Education Review*, 40, 139-157.
- Peterson, P. L., Carpenter, T., & Fennema, E. (1989). Teachers' knowledge of students' knowledge in mathematics problem solving: Correlational and case analyses. *Journal of Educational Psychology*, 81, 558-569.

- Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, 38, 109–119.
- Pianta, R.C., LaParo, K.M., & Hamre, B.K. (2008). *Classroom Assessment Scoring System Manual, K-3*. Baltimore, MD. Paul H. Brookes Publishing Co.
- Post, T., Harel, G., Behr, M., & Lesh, R. (1988). Intermediate teachers' knowledge of rational number concept. In Fennema, et al. (Eds.), *Papers from First Wisconsin Symposium for Research on Teaching and Learning Mathematics* (pp.194-219). Madison, WI: Wisconsin Center for Education Research.
- Putnam, R. T., Heaton, R. M., Prawat, R. S., & Remillard, J. (1992). Teaching mathematics for understanding: Discussing case studies of four fifth-grade teachers. *The Elementary School Journal*, 93, 213-228.
- Raymond, A. M. (1997). Inconsistency between a beginning elementary school teacher's mathematics beliefs and teaching practice. *Journal for Research in Mathematics Education*, 28, 550-576.
- Raudenbush S.W., & Bryk, A.S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. 2nd ed. Thousand Oaks, CA: Sage Publications.
- Raudenbush, S., Bryk, A., Cheong Y. F., Congdon, R., & Toit, M. (2004). *HLM 6: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.
- Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from Panel data. *The American Economic Review*, 94, 247-252.
- Rowan, B., Chiang, F-S., & Miller, R. J. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education*, 70, 256-284.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15, 4-14.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57, 1- 22.

- Silver, E. A., & Stein, M. K. (1996). The quasar project: The “revolution of the possible” in mathematics instructional reform in urban middle schools. *Urban Education, 30*, 476-521.
- Simon, M. A. (1993). Prospective elementary teachers’ knowledge of division. *Journal of Research in Mathematics Education, 24*, 233-254.
- Sloane, F. (2003). *An assessment of Sorensen's model of school differentiation: A multilevel model of tracking in middle and high school mathematics* (Unpublished doctoral dissertation). The University of Chicago, Illinois.
- Sloane, F. (2008a). Comments on Slavin: Through the looking glass: Experiments, quasi-experiments, and the medical model. *Educational Researcher, 37*, 41-46.
- Sloane, F. (2008b). Randomized trials in mathematics education: Recalibrating the proposed high watermark. *Educational Researcher, 37*, 624-630.
- Sloane, F., Holding, B., & Kelly, A. E. (2008). Longitudinal analysis and interrupted time series designs: Opportunities for the practice of design research. In A. E. Kelly, R. Lesh, and J. Baek (Eds.), *Handbook of design research in education: Innovations in science, technology, mathematics and engineering* (pp. 449-458). New York: Routledge.
- Stein, M. K., Baxter, J. A., & Leinhardt, G. (1990). Subject-matter knowledge and elementary instruction: A case from functions and graphing. *American Educational Research Journal, 27*, 639-663.
- Stein, M.K., Smith, M.S., Henningsen, M.A. & Silver, E.A. (2000). *Implementing standards-based mathematics instruction*. New York: Teachers College Press.
- Stigler, J., & Hiebert, J. (1999). *The teaching gap*. New York: Free Press
- Tirosh, D., & Graeber, A. O. (1990). Cognitive conflict to explore preservice teachers’ thinking about division. *Journal for Research in Mathematics Education, 21*, 98-108.
- Wenglinsky, H. (2002, February 13). How schools matter: The link between teacher classroom practices and student academic performance. *Education Policy Analysis Archives, 10*. Retrieved from <http://epaa.asu.edu/epaa/v10n12/>.

- Williams, S., & Baxter, J. (1996). Dilemmas of discourse-oriented teaching in one middle school mathematics classroom. *Elementary School Journal*, 97, 21-38.
- Willett, J. (1988). Questions and answers in the measurement of change. In E.Z. Rothkopf (Ed.), *Review of research in education* (pp. 345-422). Washington, DC: AERA.
- Willett, J. (1989). Some results on reliability for the longitudinal measurement of change: Implications for the design of studies of individual growth. *Educational And Psychological Measurement*, 49, 587-602.
- Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, 11, 57-67.
- Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*, 27, 458-477.

APPENDIX A

UNIVERSITY HUMAN SUBJECTS INSTITUTIONAL REVIEW BOARD



Research Compliance Office
Office for Research & Sponsored Projects Administration
P.O. Box 873503
Tempe, AZ 85287-3503

Phone
(480) 965-6788
Facsimile
(480) 965-7772

To: Finbarr Sloane

From: Mark Roosa, Chair
Institutional Review Board

Date: 09/18/2006

Committee Action: **Exemption Granted**

IRB Action Date: 09/18/2006

IRB Protocol #: 0608001014

Study Title: Moving Teachers and Students from Arithmetic to Algebraic Thinking in Murphy School District (K-5 Insight)

The above-referenced protocol is considered exempt after review by the Institutional Review Board pursuant to Federal regulations, 45 CFR Part 46.101(b)(1) .

This part of the federal regulations requires that the information be recorded by investigators in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects. It is necessary that the information obtained not be such that if disclosed outside the research, it could reasonably place the subjects at risk of criminal or civil liability, or be damaging to the subjects' financial standing, employability, or reputation.

APPENDIX B

SAMPLE DAP TEST ITEMS

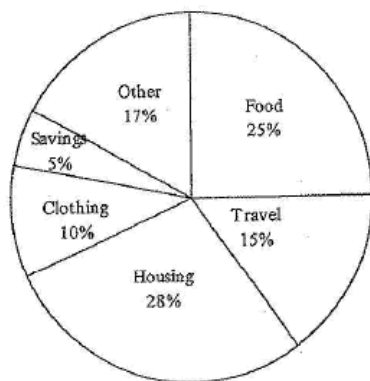
Strand 1: Number and Operations

1 The fraction $\frac{2}{5}$ is equivalent to ____.

- A .10
- B .25
- C .40
- D 2.5

Strand 2: Data Analysis, Probability, and Discrete Mathematics

27 Melissa earns \$2000 a month as an office assistant. The circle graph shows how she spends her money each month.



What does she spend most of her money on each month?

- A Food and housing
- B Travel and housing
- C Housing and travel
- D Clothing and housing

Strand 3: Patterns, Algebra, and Functions

39 Look at the number pattern below.

15, 17, 34, 36, __, __

Which of the following numbers extends the pattern?

- A 72, 74
- B 75, 120
- C 120, 220
- D 120, 240

Strand 4: Geometry and Measurement

53 Which figure below is an example of an acute angle?

A



B



C



D



Strand 5: Structure and Logic

- 67 Kelly wants to buy three new DVDs that cost \$14.50 each. In two weeks, she saved \$18.40 from her allowance. How many weeks will she have to save her entire allowance before she can buy the movies?

What relevant information is missing to solve this problem?

- A how many movies she already has
- B the number of weeks she's wanted the movies
- C that the movies were marked down from \$15.25
- D the allowance amount she gets each week.